

LEXICAL SIMPLIFICATION AND ELABORATION:  
SENTENCE COMPREHENSION  
AND INCIDENTAL VOCABULARY ACQUISITION

A THESIS SUBMITTED TO THE GRADUATE DIVISION OF THE  
UNIVERSITY OF HAWAII IN PARTIAL FULFILLMENT  
OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS

IN

ENGLISH AS A SECOND LANGUAGE

DECEMBER 2000

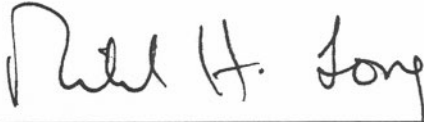
By  
Ken Urano

Thesis Committee

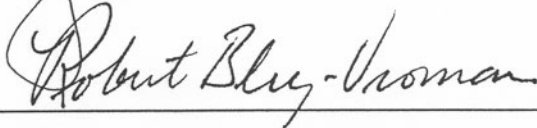
Michael H. Long, Chairperson  
Robert Bley-Vroman  
James Dean Brown

We certify that we have read this thesis and that, in our opinion, it is satisfactory in scope and quality as a thesis for the degree of Master of Arts in English as a Second Language.

THESIS COMMITTEE

  
Chairperson





© Copyright 2000

by

Ken Urano

## ACKNOWLEDGMENTS

I wish to express my sincere gratitude to my committee members, Professors Michael H. Long, Robert Bley-Vroman, and J. D. Brown, for their continuous support and encouragement as well as their invaluable suggestions and feedback throughout the period I worked on my thesis. I also wish to thank Professor Craig Chaudron, who kindly agreed to give me support as a committee member while Professors Long and Brown were not in Hawai'i.

My thanks also go to the participants of my study, many of whom are good friends of mine. My M.A. thesis would never have been completed without their participation.

I would also like to thank Cathy Harrison for letting me use her (or her husband's) computer for the data collection.

Last, but not at all least, I would wholeheartedly like to express my gratitude to my parents for their understanding and support.

## ABSTRACT

The present study investigated the effects of two different types of input modification, simplification and elaboration, on second language comprehension and incidental vocabulary acquisition. 40 Japanese learners of English as a second language were asked to read English sentences with different types of modification: 10 unmodified baseline sentences, 10 lexically simplified sentences, 10 lexically elaborated sentences, and 10 distractors. Reading comprehension was measured by mean reading times and comprehension questions on each sentence. Right after the treatment, two vocabulary tests, i.e., a form-recognition test and a meaning-recognition test, were used to measure incidental vocabulary acquisition. Results suggested that (a) both lexical simplification and elaboration can improve learner comprehension at the sentence level; (b) lexical elaboration triggers incidental vocabulary acquisition, while simplification does not; and (c) learners of higher proficiency benefit more from lexical elaboration in terms of the acquisition of word meanings.

## TABLE OF CONTENTS

ACKNOWLEDGMENTS .....	iv
ABSTRACT.....	v
TABLE OF CONTENTS .....	vi
LIST OF TABLES .....	viii
LIST OF FIGURES .....	ix
CHAPTER 1: INTRODUCTION.....	1
CHAPTER 2: REVIEW OF THE LITERATURE.....	3
Target and type of modification.....	3
Simplification, elaboration, and second language comprehension.....	6
Problems and limitations of previous studies .....	12
Lexical simplification and elaboration.....	15
Purpose of the study.....	16
CHAPTER 3: METHOD .....	18
Participants.....	18
Materials .....	18
Test sentences.....	18
Vocabulary tests .....	24
Procedures.....	24
Design and analyses.....	26
Hypotheses.....	29

CHAPTER 4: RESULTS .....	32
Sentence comprehension measures.....	32
Mean reading times.....	32
Comprehension questions .....	34
Vocabulary acquisition measures.....	42
Form-recognition test.....	42
Meaning-recognition test .....	44
CHAPTER 5: DISCUSSION.....	49
Hypotheses.....	49
Lexical simplification, elaboration, and sentence comprehension .....	50
Lexical simplification, elaboration, and incidental vocabulary acquisition .....	52
CHAPTER 6: CONCLUSION .....	60
Summary of the findings and implications .....	60
Limitations of the study and suggestions for future studies .....	61
NOTES.....	63
APPENDIX A: TEST SENTENCES AND COMPREHENSION QUESTIONS .....	64
APPENDIX B: FORM-RECOGNITION TEST .....	70
APPENDIX C: MEANING-RECOGNITION TEST .....	71
APPENDIX D: INSTRUCTIONS FOR THE EXPERIMENT .....	73
REFERENCES .....	75

## LIST OF TABLES

<u>Table</u>	<u>Page</u>
1. Summary of the Studies on Simplification, Elaboration, and L2 Comprehension .....	10
2. Target Words, Distractors, and High-frequency Synonyms.....	21
3. Mean Reading Times and Standard Deviations .....	32
4. Means and Standard Deviations for Comprehension Scores .....	34
5. MANOVA Table for Comprehension Measures .....	37
6. Two-way Repeated-measures ANOVA Table for Mean Reading Times.....	39
7. Two-way Repeated-measures ANOVA Table for Comprehension Scores.....	41
8. Form-recognition Test Mean and Standard Deviations.....	42
9. Meaning-recognition Test Means and Standard Deviations .....	44
10. MANOVA Table for Vocabulary Tests.....	46
11. One-way Repeated-measures ANOVA Table for the Form-recognition Test .....	47
12. Post Hoc Comparisons Table .....	47
13. One-way Repeated-measures ANOVA Table for the Meaning-recognition Test.....	48
14. Means and Standard Deviations for Meaning-recognition Test Scores for Two Proficiency Levels.....	55
15. Two-way Repeated-measures ANOVA Table for Meaning-recognition Test for Two Proficiency Levels.....	57
16. Post Hoc Comparisons Table: High-proficiency Group .....	58
17. Post Hoc Comparisons Table: Low-proficiency Group.....	58

## LIST OF FIGURES

<u>Figure</u>	<u>Page</u>
1. Two Independent Variables .....	20
2. Test Forms.....	27
3. Mean Reading Times .....	33
4. Mean Comprehension Scores .....	35
5. Form-recognition Test Mean Scores .....	43
6. Meaning-recognition Test Mean Scores .....	45
7. Mean Meaning-recognition Test Scores for Two Proficiency Levels.....	56

# CHAPTER 1

## INTRODUCTION

It has been widely accepted among second language acquisition (SLA) researchers that exposure to the target language input is necessary for SLA. Moreover, input needs to be comprehended by second language (L2) learners for subsequent language acquisition processes to take place (Chaudron, 1985). Therefore, it is of theoretical and empirical interest to inquire how input is made comprehensible to learners with limited L2 proficiency.

Motivated by studies of caretaker speech in first language acquisition (e.g., Snow, 1972), SLA researchers have investigated native speakers' (NSs') input adjustments for L2 learners. Input addressed to L2 learners is often called *foreigner talk* (Ferguson, 1975), or *teacher talk* in the classroom situation (Chaudron, 1988), and researchers have discovered that input to non-native speakers (NNSs) is modified in various ways, in both the spoken and written mode.

Recent research on input modification addresses the relative effectiveness of two different types of modification, i.e., simplification and elaboration, on L2 comprehension. It is commonly believed that simplifying input, or manipulating linguistic items unknown to the learner, will enhance L2 comprehension, and many commercially published L2 materials have applied this technique (Yano, Long, & Ross, 1994). On the other hand, several researchers have argued against the use of simplification because (a) simplifying input does not necessarily aid comprehension (e.g., Blau, 1982) and (b) it removes from the input linguistic items that L2 learners need to learn (Long, 1983; 1996; Yano et al., 1994).

Although previous studies have generally found that both simplification and elaboration improve L2 comprehension, we are still in need of further research in order to confirm their relative effectiveness. Moreover, few studies have explored the effects of simplification and elaboration on learning of new linguistic items. The present study intends to investigate the effects of two types of modification, i.e., simplification and elaboration, on both L2 comprehension and SLA, with a focus on the lexical domain in English.

## CHAPTER 2

### REVIEW OF THE LITERATURE

#### *Target and type of modification*

When one discusses the role of input modification, it is useful to bear two questions in mind: (a) What is modified? (target of modification); and (b) How is it modified? (type of modification). Since the 1970's, a number of studies have described the characteristics of input modification by comparing NSs' speech to NNSs with their talk to other NSs or NNSs of higher proficiency.

With regard to the first question, *What is modified*, researchers have investigated modifications at different linguistic levels, i.e., phonology, lexis, syntax, and discourse (for comprehensive reviews of the literature, see Chaudron, 1988, Chapter 3; and Long, 1996). In the phonological domain, studies suggest that NSs speak more slowly, speak with clearer articulation, and insert more pauses, when addressing NNSs than when they speak to NSs (Ishiguro, 1986; Kleifgen, 1985; Wesche & Ready, 1985). For example, in their comparative study of university classroom discourse, Wesche and Ready (1985) found that an English-speaking professor spoke to L2 students significantly slower than to L1 students (1,262 and 1,883 words, respectively, in 14-minute transcripts).<sup>1</sup> The professor was found to pause significantly more frequently when speaking to L2 students than to L1 students (e.g., 28 and eight pauses of two or more seconds).

At the lexical level, Chaudron (1982), among others, reports that NS teachers in four different school systems in Canada used high-frequency, "basic" vocabulary more often with ESL learners (NNSs) than with other NS students.

In terms of modification of syntax, several studies found that NSs speak to NNSs in significantly shorter utterances than they use with NSs (e.g., Wesche & Ready, 1985). Other researchers report that NSs modify utterance length according to the different proficiency levels of NNS addressees (Kleifgen, 1985; Ishiguro, 1986).

NS speech modification at the discourse level is also referred to as *interactional modification*, as opposed to the other types of modification mentioned above, which are categorized as *linguistic modification* (Long, 1981, 1983, 1996). NS's use of self-repetition is considered an instance of modification at the discourse level. Ellis (1985), for example, counted the same NS teacher's use of self-repetition on two occasions in classroom interaction with two adolescent ESL students, and found that the frequency of self-repetition decreased significantly over the six-month period of observation. He concluded that this decrease was due to the learners' improved L2 comprehension, which suggests that more self-repetition will be used with less proficient learners.

To the second question, *How is input modified?*, researchers suggest that there are two different types of modification, i.e., simplification and elaboration. Simplification can be defined as controlling the text targeted at L2 learners by removing unfamiliar linguistic items (e.g., unknown grammatical constructions and lexis) in order to enhance comprehension. It has been widely applied in many commercially published L2 reading materials in the belief that the use of controlled vocabularies and short, simple sentences will facilitate L2 reading comprehension (Yano et al., 1994). Most of the readability formulas are also based on lexical and syntactic complexity, such as the ratio of low and high frequency words, or sentence length (Baker, Atwood, & Duffy, 1988). Therefore, simplified texts are by definition considered easier to understand than unmodified texts if

those formulas are used as measures.

Although simplification is generally considered a good candidate for a text modification technique for improving comprehension, several researchers argue against its use. In first language (L1) reading research, Green and Olsen (1988) found that readability-adapted (i.e., simplified) materials were not significantly easier for children to understand than the original texts. In an L2 reading study, Blau (1982) demonstrated that simple sentences alone do not necessarily aid comprehension (this issue will be discussed in more detail later).

In terms of language learning, even if simplification may facilitate learner comprehension, it has a crucial weakness in that comprehension is achieved by removing items that learners need to learn (Long, 1983, 1996; Long & Ross, 1993; Yano et al., 1994). To quote Yano et al. (1994):

. . . removal of possibly unknown linguistic items from a text may facilitate comprehension but will simultaneously deny learners access to the items they need to learn. Linguistic simplification can be self-defeating to the extent that the purpose of a reading lesson is not the comprehension of a particular text, which learners are unlikely ever to encounter again outside the classroom, but the learning of the language in which the text is written and/or the development of transferable, non-text-specific, reading skills. (p. 191)

An alternative text-modification technique is *elaboration*. A text can be modified for easier comprehension not by removing complex structures, as simplification does, but by

adding redundant information to the text through the use of repetition, paraphrases, and appositionals (Long, 1996, p. 422).

As an example of each modification technique at the lexical level, consider 1 - 3 below from Urano (1998):

- (1) Baseline version: Everybody knows that Ken is *diligent* and kind to others.
- (2) Simplified version: Everybody knows that Ken is hardworking and kind to others.
- (3) Elaborated version: Everybody knows that Ken is *diligent*, or hardworking, and kind to others.

Assuming that the word *diligent* in Sentence 1 is unknown to the learner, the simplified version (2) and the elaborated version (3) facilitate learner comprehension in different ways. In Sentence 2, the target word *diligent* is replaced by a high-frequency, supposedly easier, synonym, *hardworking*. On the other hand, the elaborated version (3) keeps the target word and then adds the synonym in apposition to the target word. If both simplification and elaboration facilitate comprehension, and if one accepts the above argument against simplification in terms of language learning, then the elaborated version is preferable since it facilitates learner comprehension of the sentence while at the same time preserving an opportunity to learn the unknown vocabulary item, *diligent*. To support this argument, however, it is necessary to demonstrate that simplification and elaboration *do* facilitate L2 comprehension.

### ***Simplification, elaboration, and second language comprehension***

Since the early 80's, a number of studies have examined the effects of different types

of text modification on L2 comprehension (for reviews, see Parker & Chaudron, 1987; Yano et al., 1994; and especially, Chung, 1995). There have been four studies that directly compared the effects of simplification and elaboration on L2 comprehension (Brown, 1985; Tsang, 1987; Yano et al., 1994; and Chung, 1995). These four studies are first reviewed in this section, then other relevant studies will be introduced.

Brown (1985) tested the reading comprehension of 30 ESL students in Taiwan with three different versions of a text as the independent variable: the NS baseline version at the 10th grade level of difficulty; the Modified Input version (i.e., simplified version) with modification in sentence structure and vocabulary; and the Modified Interactional Structure version (i.e., elaborated version), which retained the native speaker level of difficulty, but elaborated information through redundancy and repetition. The results of the 20 multiple-choice questions after reading the text showed that scores of both simplified and elaborated conditions were significantly higher than that of the NS baseline condition. Although the elaborated group obtained a higher score than the simplified group, the difference was not statistically significant.

In a replication study of Brown (1985), Tsang (1987) found a similar tendency with 401 ESL students in Hong Kong. Using the same texts as Brown (1985), she examined the differences in comprehension among the baseline, simplified, and elaborated texts with Cantonese-speaking students at five different grade levels (Grades 9 through 13). The results for the lowest two grades (i.e., Grades 9 and 10) displayed significantly higher scores for the simplified and elaborated groups compared with the baseline group.

Yano et al. (1994; also, Ross, Long, & Yano, 1991) advocated use of elaboration as an alternative to simplification. As noted earlier, they argued against the use of simplification

since it deprives the L2 learner of an opportunity to learn new linguistic items. Thirteen reading passages with three different forms (i.e., NS baseline, simplified, and elaborated) were prepared for their experiment with 483 Japanese college students. Simplified texts were created by keeping the length of sentences and the number of multisyllabic words and embedded clauses to a minimum. Elaborated versions were made by parenthetical paraphrase or providing definitions of low-frequency content words in the original. The scores of 30 comprehension questions showed that the simplified version was significantly easier to comprehend than the baseline condition, but no statistically significant difference was found between the elaborated condition and the baseline condition. Yano et al. further analyzed three subcategories of question types, and found that elaboration outperformed both simplified and baseline versions in mean scores of what they termed "inference" items. However, this result should be interpreted with caution because of the small number of items ( $k = 2$ ).

After reviewing previous studies, Chung (1995) hypothesized that both simplification and elaboration would facilitate L2 reading comprehension. She was also interested in the effects of different subtypes of elaboration, i.e., lexical elaboration and structural elaboration. Chung prepared five different versions of a reading passage: unmodified baseline, simplified, lexically elaborated, structurally elaborated, and lexically & structurally elaborated. In the simplified version, Chung rewrote compound sentences as several simple declarative sentences which expressed a single main idea per sentence, used the active voice where possible, and substituted low-frequency words with high-frequency ones. Lexical elaboration was achieved by adding redundancy, e.g., definition, synonym, and hyperonym, to the language items anticipated to be unknown to the participants.

Similarly, structural elaboration was achieved by adding redundancy to the text in order to clarify message content and organization through signaling of intersentential relationships, retention of full NPs, supplying omitted elements, using anaphoric rather than cataphoric reference, and paraphrasing for summary statements which make already existing logical relations explicit without adding new information. The lexically & structurally elaborated text was generated by combining lexical and structural elaboration. The results of the 20-item multiple-choice comprehension test showed a significant difference between the baseline condition and the simplified condition. On the other hand, no significant differences were found between the elaborated conditions and the baseline condition.

Table 1 summarizes the results of the studies reviewed above. First of all, all four studies found significant effects of simplification on L2 comprehension. The effects of elaboration, on the other hand, are not so consistent as those of simplification, although the difference in comprehensibility between elaboration and simplification seems to be small.

**Table 1***Summary of the Studies on Simplification, Elaboration, and L2 Comprehension*

Study					
Brown (1985)	elaborated	>ns	simplified	>*	baseline
	elaborated			>*	baseline
Tsang (1987)	simplified	>ns	elaborated	>*	baseline
	simplified			>*	baseline
Yano et al. (1994)	simplified	>ns	elaborated	>ns	baseline
	simplified			>*	baseline
(inference items)	elaborated	>*	simplified	>*	baseline
	elaborated			>*	baseline
Chung (1995)	simplified	>ns	elaborated	>ns	baseline
	simplified			>*	baseline

*Note.* >\* = statistically significant difference, >ns = non-significant difference

Other studies which examined the effects of a single type of modification generally show similar results. However, there are exceptions. Blau (1982) created three different versions of a text: Version 1 with short, simple sentences; Version 2 with complex sentences with clues to underlying relationships; and Version 3 made up of complex sentences without clues to underlying relationships. Scores on 24 multiple-choice comprehension questions by two groups, i.e., ESL university students and ESL 8th graders, showed no significant difference. Blau argued that use of simple sentences, one of the main criteria for common readability formulas, does not necessarily aid comprehension.

Parker and Chaudron (1987) defined elaboration in a slightly different way from others. They categorized modification into three types: two types of modifications of input (simplification & elaboration), and modification of interaction. They operationally defined elaboration as "the addition of redundancy, and the explicit realization of underlying thematic relations" (p. 110). Two different types of passages were created from a passage for their study. An elaborated version retained all the redundancy and thematic structure, and was modified with some additional changes of the same type. The other text, the non-elaborated version, had all redundancies eliminated, and all the thematic structure reduced to canonical word order form. Forty-three college students answered a reading comprehension cloze test after reading one of two different types of passages. No statistically significant difference was found between the scores of the unmodified version and the elaborated version. In her review of the literature, Chung (1995) pointed out this null finding may have been due to the fact that even the non-elaborated version was actually *modified* in that all redundancies were eliminated and all the thematic structure was reduced to canonical word order. In other words, the characteristics of the

non-elaborated text were similar to those of *simplified* texts. Therefore, what Parker and Chaudron (1987) studied was not necessarily the effects of elaboration relative to the NS baseline text, but rather a comparison of elaboration with simplification. In order to investigate the effects of elaboration, the use of *unmodified* baseline text was needed in the control condition.

### ***Problems and limitations of previous studies***

From the review of the literature, it can be stated that, generally speaking, text modifications seem to have positive effects on L2 comprehension. However, it is obvious that more research is needed to better understand the role of different kinds of modification. The results of previous studies have not yet shown that particular types of modifications are consistently effective for L2 comprehension. As Yano et al. (1994) note, we are still in need of "a larger, more carefully controlled study of the relative effectiveness of simplification and elaboration" (p. 200).

There are a few potential limitations in the previous studies that may have caused the variation of their findings. First, the measures of L2 comprehension used may not have directly assessed *comprehension*. Most of the studies, except those of interactional modifications (e.g., Pica, Doughty, & Young, 1986; Pica, Young, & Doughty, 1987), employed paper-and-pencil comprehension tests, such as multiple-choice questions, recall, and/or cloze, *after* the participants read/listened to the test passages. These measures may leave room for other intervening variables, e.g., participants' limited memory, that might have made the results less clear.

Another possible reason for the less consistent findings is that the comprehension tests

employed a nominal scale (i.e., correct/wrong), which may not be sensitive enough to distinguish the effects of different types of modification. The use of a nominal scale requires participants clearly to understand (part of) the passage they read/listen to in order to demonstrate comprehension. In other words, it may be the case that the results of previous studies are less straightforward not because text modifications did not facilitate comprehension, but because their effects were so subtle that they could not be detected by the comprehension measures employed.

Third, the amount, or frequency, of modifications contained in the passages was not controlled in previous studies. If text modification helps L2 comprehension, the amount of modifications in a passage may influence the comprehensibility of the text. Consider the following case in which a simplified version and an elaborated version are generated from the same baseline text. If, for example, 10 words are simplified in the simplified version and 100 words are elaborated in the elaborated version, the results of the comprehension test would probably favor the elaborated version simply because the elaborated version contains more frequent modifications than the simplified version. Without controlling this variable, it would be difficult to compare different types of modification.

Lastly, although the ultimate goal of the studies of text modification is to investigate the effects of modification on second language acquisition, all the studies except Chung (1995) stopped at L2 comprehension and did not examine the role of modification in learning. Chung (1995) utilized two different types of vocabulary test, a form-recognition test and a meaning-recognition test, to assess incidental vocabulary acquisition after reading the two different versions of the same text: unmodified baseline, simplified, lexically elaborated, structurally elaborated, and lexically & structurally elaborated. Even

though she found significant effects on L2 comprehension for the simplified text, as reported earlier, no significant difference was found among different modification types in either of the two vocabulary tests. Since Chung's study is open to the same criticisms discussed in this section, whether or not modifications facilitate L2 vocabulary acquisition cannot be determined.

In order to solve, or at least minimize, the problems of previous studies, two revisions are proposed in the present research. First, the use of reaction time as a measure of L2 comprehension would possibly identify subtle differences in effects for different types of modification. Reaction time (RT) has been developed as a supplemental measure to grammaticality judgment used in psycholinguistic studies within the UG paradigm (Bley-Vroman & Masterson, 1989). Freedman and Foster (1985) set up a sentence-matching task (i.e., a task to find out if two sentences are identical or not), and demonstrated that the matching task took longer when the sentences were ungrammatical. RT is also used in L1 reading studies, there referred to as reading time. Kemper (1988), for example, found that the mean reading times for target sentences were significantly longer with passages in which some sentences referring to critical actions, physical states, or mental states were deleted, which suggests that those sentences were more difficult to understand.

It is assumed in a text modification study that a mean reading time becomes shorter when a certain type of modification facilitates L2 comprehension. Since this measure is a ratio, rather than a nominal scale (i.e., correct/wrong), like other measures used in previous studies, it is expected to successfully distinguish levels of comprehensibility. However, this method has a limitation, too. Since RT is a sensitive measure, careful attention should

be paid to control all factors other than text modification. Testing L2 comprehension at the sentence level, rather than over the whole passage, will be helpful. By comparing modified versions of the same sentence, the influence of unwanted variables, such as different lexical items used in the sentence, can be minimized, if not completely eliminated.

Second, effects of simplification and elaboration on both L2 comprehension and acquisition need to be tested in one study. As discussed, there has been only one study, Chung (1995), that has attempted to do this, and the results were not conclusive.

### ***Lexical simplification and elaboration***

Lexical simplification and elaboration was chosen for this study for the following two reasons. First, several studies have investigated the relationship between text modifications and vocabulary acquisition (e.g., Chung, 1995; Kim, 1996; Watanabe, 1992). Although these researchers had slightly different interests, they all tested the effects of text modification on vocabulary acquisition as part of their studies. Although the results generally suggested positive effects of text modification on vocabulary acquisition, a significant difference was found only by Kim (1996). Thus, there is still scope for a new study in this area. By focusing on modification and acquisition of lexical items, a clear causal relationship between modification and comprehension/acquisition can be examined.

The second reason is a practical one. It is difficult to test the effects of modification of other linguistic items, such as syntax, at the sentence level. It seems impossible to create a comparable syntactically simplified and elaborated sentence pair. Sentences can be syntactically simplified in some cases, e.g., by using canonical word order, and then the comprehensibility of simplified sentences can be compared with that of the baseline NS

sentences (e.g., Issidorides & Hulstijn, 1992). However, there seems to be no syntactically elaborated sentences that can be compared with the simplified version, since syntactical elaboration often involves modification at the intersentential level, such as repetition or addition of a clause, or paraphrasing a whole sentence, which makes sentence-level comparisons difficult. On the other hand, sentences can be lexically simplified and elaborated with relative ease. In a simplified version, vocabulary items unknown to the learner can be replaced by high-frequency synonyms, for example. To elaborate a sentence lexically, high-frequency synonyms can be attached in apposition to the unknown words.

Lexical simplification in this study is more specific than that in earlier research (e.g., Parker & Chaudron, 1987). It is defined as substituting unknown words with high-frequency "basic" vocabulary items. Lexical elaboration, on the other hand, is achieved not by deleting those unknown words, but by adding their synonyms in apposition to them.

### ***Purpose of the study***

The purpose of this study is to investigate the effects of lexical simplification and elaboration on L2 sentence comprehension and incidental vocabulary acquisition. This will be achieved by (a) measuring the mean reading time of sentences with different modifications; and (b) vocabulary tests administered after (a). The following research questions are addressed:

1. To what degree, if any, do lexical simplification, elaboration, or both, facilitate L2 sentence comprehension?

2. To what degree, if any, do lexical simplification, elaboration, or both, influence incidental vocabulary acquisition, either positively or negatively?

## CHAPTER 3

### METHOD

#### *Participants*

Forty native speakers of Japanese, 33 female and seven male, participated in the study on a voluntary basis. All of them were enrolled in intensive English language programs in the State of Hawai'i at the time of their participation. Their mean age was 27.10 years, with a range of 18 to 49. Twenty-two of them had taken TOEFL at least once<sup>2</sup>, and their most recent scores ranged from 350 to 573 with a mean of 466.46. They had all experienced at least six years of formal English language education in Japan at the secondary school level. Their length of stay in English-speaking countries ranged from one month to ten years, with a mean of 11.00 months. Each participant received a free movie ticket or a debit card for photocopying at the university libraries as compensation after the experiment.

#### *Materials*

*Test sentences.* In the previous chapter, lexical simplification and elaboration were operationally defined as substituting unknown words with high-frequency synonyms, and as adding synonyms of the unknown words in apposition to them, respectively. Sentences 4 - 6 are samples from the actual test sentences used in the experiment.

(4) Baseline version: Becky could not eat the cake because it was too *vight* for her.

(5) Simplified version: Becky could not eat the cake because it was too sweet for her.

(6) Elaborated version: Becky could not eat the cake because it was too *vight*, or sweet, for her.

In this example, the nonsense word *vight* is the target item. It is synonymous with a high-frequency word, *sweet*. In the simplified condition (5), the target word is replaced by an easier synonym. The elaborated sentence (6) retains the target word, and the synonym in the appositional position gives participants a clue as to its meaning.

As can be seen from the above three sentences, elaborated sentences are always two words longer than the other types of sentences. Since mean reading times were used as a measure of comprehension, this difference in sentence length could negatively influence the results. It can be assumed that reading times become longer with longer sentences. A possible solution to deal with this problem is to adjust the mean reading time by the length of the sentence (e.g., number of words in the sentence), as in Urano (1998). However, it is not clear if there is a simple linear relationship between the length of a sentence and reading time, which is the presupposition of this type of adjustment (R. Bley-Vroman, personal communication, November 18, 1998).

To deal with this problem, another version of the sentence was prepared so that the difference in length can be statistically treated (7).

(7) Distractor version: Becky could not eat the cake because it was too *vight*, or malash, for her.

This sentence is the same as the elaborated version in sentence length. However, another nonsense word, *malash*, is added to the original. Addition of a nonsense word will not help participants comprehend the meaning of the sentence. With this Distractor Version in it, a

two-way repeated-measures design was planned (Figure 1).

		sentence length	
		+ (long)	- (short)
high-frequency	+	elaborated	simplified
synonym	-	distractor	baseline

**Figure 1. Two Independent Variables**

In a pilot for the present study, Urano (1998) defined an *unknown* vocabulary item as a very low-frequency word which appears only once or twice in Dahl (1979), a frequency list of words in spoken American English. However, the results of one of the vocabulary tests suggested that, in fact, some of the participants already knew some of those target words before the experiment took place. Therefore, nonsense words were created and used in this study to ensure that they were all unknown to the participants.

Eighty nonsense words were first created by changing one phoneme of existing English words, while obeying the morpho-phonetic rules of English. Those nonsense words were then randomly assigned to one of the following two groups: target words and distractors. Next, forty content words were selected as high-frequency words from the 2000-word Longman Defining Vocabulary (Longman, 1987). Each target word was then randomly matched with a distractor and a high-frequency word. (See Table 2 for a list of 40 sets of the target words, distractors, and high-frequency synonyms.)

**Table 2*****Target Words, Distractors, and High-frequency Synonyms*****(in order of appearance in the experiment)**

Target Words	Distractors	High-frequency Synonyms
( <i>k</i> = 40)	( <i>k</i> = 40)	( <i>k</i> = 40)
plear	hister	height
darm	splim	wet
ober	hinch	bag
smub (smubbing)	initate (initating)	cry (crying)
ungle	vonce	buy
rascine	ruspence	bird
litidate	feach	kind
rimmor	tiver	stop
praverse	torth	notice
mosat	armot	sharp
serald	cratic	accident
marfle	golt	quiet
butty	lipal	bread
exlume	tostle	dirty
nace	soor	entrance
selest	telt	hairy
dield	nooch	jealous

**Table 2 (Continued)***Target Words, Distractors, and High-frequency Synonyms***(in order of appearance in the experiment)**

Target Words	Distractors	High-frequency Synonyms
( <i>k</i> = 40)	( <i>k</i> = 40)	( <i>k</i> = 40)
gank	shapper	noise
cemit	finter	ocean
tepsin (tepsined)	grile (griled)	promise (promised)
sleam (sleams)	brudge (brudges)	restaurant (restaurants)
rigure	smour	rude
titch	rastle	trust
smape (smaper)	soab (soaber)	warm (warmer)
elert	trieze	allow
conpeal (conpealed)	slire (slired)	complain (complained)
mostage	anept	factory
pardel	anvoy	garage
raste	crail	persuade
ausit (ausiting)	prulic (prulicing)	cook (cooking)
pollow	fick	interesting
lassage (lassaging)	sporl (sporling)	laugh (laughing)
doral	midate	medicine
dail	benile	hat

**Table 2 (Continued)**

***Target Words, Distractors, and High-frequency Synonyms***

**(in order of appearance in the experiment)**

Target Words	Distractors	High-frequency Synonyms
( <i>k</i> = 40)	( <i>k</i> = 40)	( <i>k</i> = 40)
tundle	birb	catch
incrint	alect	funny
rebair (rebairs)	megate (megates)	poem (poems)
progue	ragget	service
ferm (fermed)	bestern (besterned)	visit (visited)
vight	malash	sweet

Forty sentences were created in such a way that one target word was included in each sentence. Each sentence consisted of between 13 and 15 words, and care was taken not to include low-frequency words in the sentences, except for the nonsense target word, to ensure that the target word would be the only word unknown to participants. After creating the sentences, the Longman list was used again to make sure that all the words except for the target words in these sentences consisted of words in the list.

Three modified versions, i.e., a simplified version, an elaborated version, and a distractor version, of each of these forty sentences were then created following the operational definitions introduced earlier. All the sentences were screened by native speakers of English for their naturalness. (See Appendix A for the complete list of the test sentences.)

***Vocabulary tests.*** Two vocabulary tests were created for the study: a form-recognition test, and a meaning-recognition test. A form-recognition test was administered since the first stage of language acquisition is considered to be recognition of the target form (Chaudron, 1985). Participants were given the forty target words printed on a sheet of paper, and were asked to circle all of those they thought they had seen in the test sentences. Ten of these target words served as distractors, since the participants did not encounter the target words in the simplified condition. The order of the forty words was randomized. (See Appendix B.)

The other vocabulary test, the meaning-recognition test, was administered in order to measure a deeper level of vocabulary acquisition. Each of the target words was followed by a list of four meanings in Japanese, the participants' L1, and they were asked to circle the one which they thought was the appropriate translation of the target word. (See Appendix C.)

### ***Procedures***

Each participant was tested individually in a quiet place without the presence of anybody else but the experimenter and themselves. The first experimental session was administered on a Macintosh PowerBook computer using PsyScope 1.2.4, a graphic user interface program designed for psychological experiments (Cohen, MacWhinney, Flatt, & Provost, 1993). This session served as the treatment (i.e., reading test sentences), and participants' comprehension was also measured by their mean reading times for each sentence.

First, a set of instructions was displayed on the screen. Although the instructions were

written in English, Japanese was used as a means of communication between the experimenter and the participants to ensure that they were fully understood. (See Appendix D for the actual instructions which appeared on the computer screen.) The participants were encouraged to ask questions before the actual test began. A given test item was presented as follows:

1. The test sentence was displayed in the center of the screen, with ordinary capitalization. The participants were asked to hit any key on the keyboard when they had finished reading it. The importance of speed, as well as understanding of the sentences, was emphasized.
2. Immediately after the participants hit any key, the test sentence was replaced by a simple yes-no question about the test sentence. Those comprehension questions were placed in order to ensure that the participants read the test sentences for meaning. It should be noted that the questions were designed so that the participants did not need to know the meaning of the target words in order to answer them correctly.
3. If the answer to the question was YES, participants were to press the *y* key; if the answer was NO, the *n* key was to be hit.
4. After the key was pressed, the words "Press any key when ready" appeared on the screen. When participants hit a key, the next test sentence appeared on the screen.

The mean reading time was measured from the onset of each test sentence until the participant presses a key to bring up the question.<sup>3</sup> A trial item was presented before the main session, and the experimenter helped the participants when they were not certain about the procedure. Also, the first item in the main session was a dummy and therefore was not included in the final analyses.

After this PsyScope session, two surprise vocabulary tests were administered to measure incidental vocabulary acquisition, first the form-recognition test, and then the meaning-recognition test.

### *Design and analyses*

The experiment was designed so that the influence of all other unwanted variables would be minimized. First of all, a full repeated-measures design was chosen to eliminate the influence of individual differences in different conditions. In other words, each participant read different sentences in all four conditions (i.e., baseline, simplified, elaborated, and distractor). Next, four test forms were created in order to counterbalance differences among test sentences, as in Figure 2. In this design, four participants could cover all 160 sentences (40 sentences x 4 types = 160 sentences). Since this experiment had 40 participants, each sentence was read by 10 participants.

	Baseline	Elaborated	Simplified	Distractor
Sentence 1	A	B	C	D
Sentence 2	D	A	B	C
Sentence 3	C	D	A	B
Sentence 4	B	C	D	A
Sentence 5	A	B	C	D
Sentence 6	D	A	B	C
Sentence 7	C	D	A	B
Sentence 8	B	C	D	A
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
Sentence 34	D	A	B	C
Sentence 35	C	D	A	B
Sentence 36	B	C	D	A
Sentence 37	A	B	C	D
Sentence 38	D	A	B	C
Sentence 39	C	D	A	B
Sentence 40	B	C	D	A

**Figure 2. Test Forms (A, B, C, & D)**

To summarize, the study involved two separate analyses: sentence-comprehension and vocabulary acquisition. The independent variables (IVs) in the sentence-comprehension test were (a) sentence length (with two levels; repeated-measures), and (b) modification in the form of provision of a high-frequency synonym (with two levels; repeated-measures). Mean reading time (in seconds) was the dependent variable (DV) in the analysis. Also, comprehension scores for the test sentences ( $k = 10$ ) were analyzed as another DV that provided supplementary information on sentence comprehension.

The other analysis, using the vocabulary acquisition tests, consisted of one IV and two DVs. Lexical modification was the independent variable (with three levels; repeated-measures), and the two vocabulary tests, i.e., the form-recognition test ( $k = 10$ ) and the meaning-recognition test ( $k = 10$ ), were the DV measures. The Distractor Condition was not included in the analyses of vocabulary acquisition since it was created to deal with the sentence length and mean reading time.

Two separate analyses of variance were conducted. One  $2 \times 2$  two-way repeated-measures multivariate analysis of variance (MANOVA) was computed, with mean reading times and comprehension scores as the dependent variables. With regard to the two vocabulary tests, a one-way repeated-measures MANOVA was used. Univariate analyses of variance (ANOVAs) were computed where significant results were found in the MANOVAs. Furthermore, in the latter analysis, post hoc multiple comparisons (Bonferroni  $t$ -tests) were computed where statistically significant differences were found in the univariate ANOVAs.

The analyses were conducted using the *SPSS Graduate Pack 9.0 for Windows* (SPSS

Inc., 1999). The significance level for the study was set at  $\alpha < .05$ . Since five separate ANOVAs were to be computed (after computing appropriate MANOVAs), the Bonferroni adjustment was used and the significance level for each analysis was  $\alpha < .01$  ( $\alpha/5 = .05/5 = .01$ ).

### ***Hypotheses***

#### *On mean reading times:*

1. There will be a significant main effect for sentence length, showing that sentence length influences mean reading time.
2. There will be a significant main effect for provision of high-frequency synonyms, showing that lexical modification (simplification and elaboration) facilitates L2 sentence comprehension.

#### *On comprehension scores:*

3. There will be no significant main effect for sentence length, showing that sentence length does not influence comprehension of the whole sentence.
4. There will be a significant main effect for provision of high-frequency synonyms, showing that lexical modification (simplification and elaboration) facilitates L2 sentence comprehension.

#### *On the form-recognition test:*

5. The mean score in the baseline condition will be significantly higher than that in the simplified condition.

6. The mean score in the elaborated condition will be significantly higher than that in the simplified condition.
7. There will be no significant difference between the means scores in the baseline and elaborated conditions.

*On the meaning-recognition test:*

8. The mean score in the elaborated condition will be significantly higher than that in the baseline condition.
9. The mean score in the elaborated condition will be significantly higher than that in the simplified condition.
10. There will be no significant difference between the mean scores in the baseline and simplified conditions.

The first hypothesis simply assumed that sentence length would influence mean reading times. The second hypothesis was motivated by the findings of previous studies. As noted earlier, both simplification and elaboration generally improve L2 comprehension. Since Kemper (1988) demonstrated that mean reading time was longer when the reader had difficulty in understanding, it was assumed that participants in this study would show the same tendency when they encounter reading difficulties. Hypothesis 3 implied that a two-word difference in length would not be an influential factor in understanding the meaning of a sentence as measured by comprehension questions. Also, it could be assumed that encountering an unknown word without any clue to its meaning would

disturb attention to the meaning of the whole sentence; thus Hypothesis 4 was formulated.

Hypotheses 5 through 7 implied that presentation of the target vocabulary would trigger the first stage of vocabulary acquisition, i.e., recognition of a word form. As Urano (1998) demonstrated, it was expected that scores on the form-recognition tests would be high in the baseline and elaborated conditions. On the other hand, no recognition should take place in the simplified condition since the target words were removed for the sake of comprehension.

Acquisition of word meanings is considered to lie at a later stage than recognition of word forms. Therefore, a difference was expected between the baseline and elaborated conditions (Hypothesis 8). Since no effort was made to help participants figure out the meaning of the target words in the baseline condition, scores on the meaning-recognition test should be lower than those in the elaborated condition, in which some hints as to the meaning of the target words, i.e., synonyms, were available. Also, since target words were not presented in the simplified condition, there should not be any learning of the vocabulary. Therefore, there would be a significant difference between the scores of the elaborated condition and the simplified condition (Hypothesis 9).

## CHAPTER 4

### RESULTS

#### *Sentence comprehension measures*

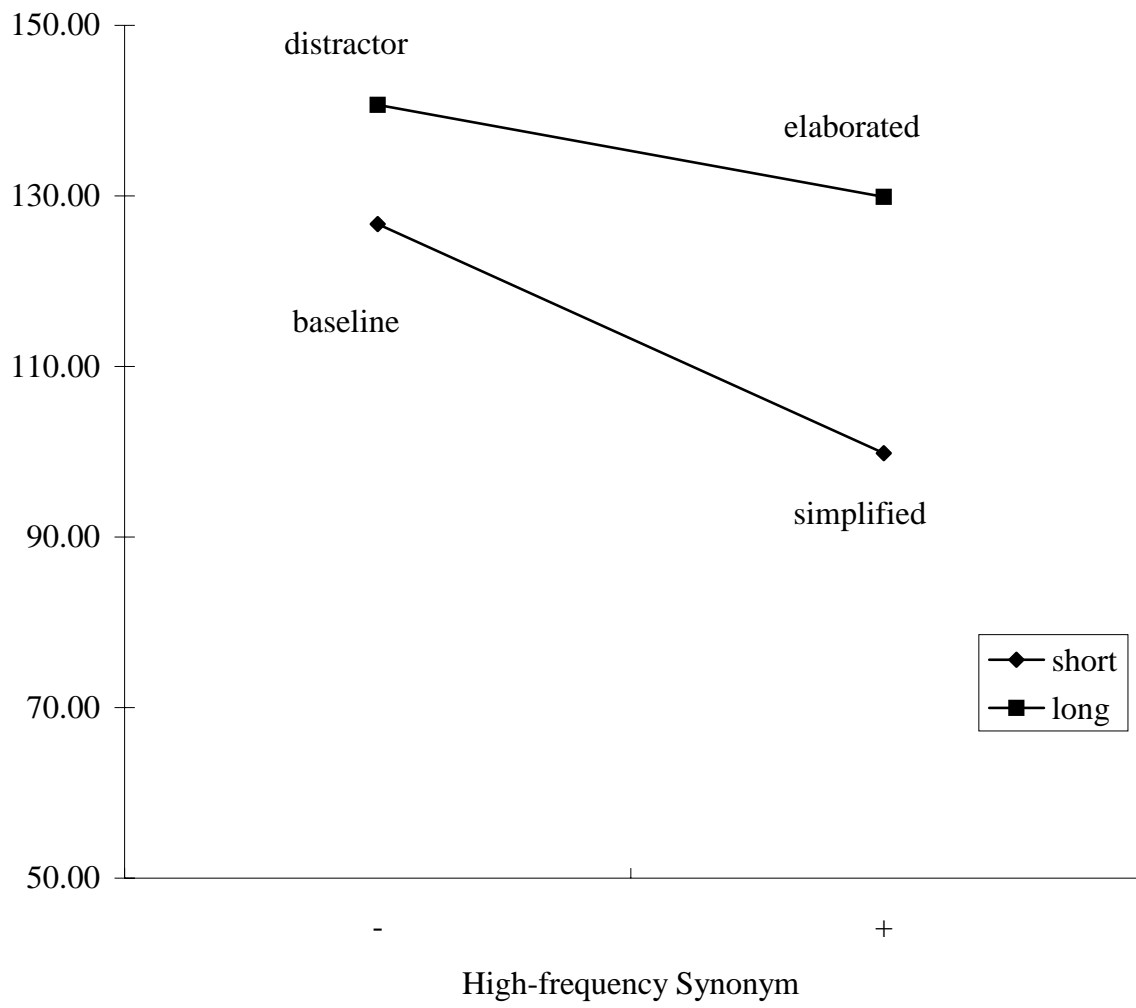
*Mean reading times.* Mean reading times and standard deviations of the four conditions are displayed in Table 3. Table 3 indicates that the mean reading time for the simplified condition (+ high-frequency synonym; - long) is the shortest and that for the distractor condition (- high-frequency synonym; + long) the longest. The mean reading times for the baseline condition (- high-frequency synonym; - long) and the elaborated condition (+ high-frequency synonym; + long) fall somewhere between these two conditions, and the difference between the baseline and elaborated conditions seems small. Figure 3 indicates the relationships between the two variables graphically.

**Table 3**

*Mean Reading Times and Standard Deviations (seconds)*

*(n = 40)*

Condition	<i>M</i>	<i>SD</i>
Baseline	126.68	36.67
Elaborated	129.90	36.08
Simplified	99.81	29.56
Distractor	140.69	34.74

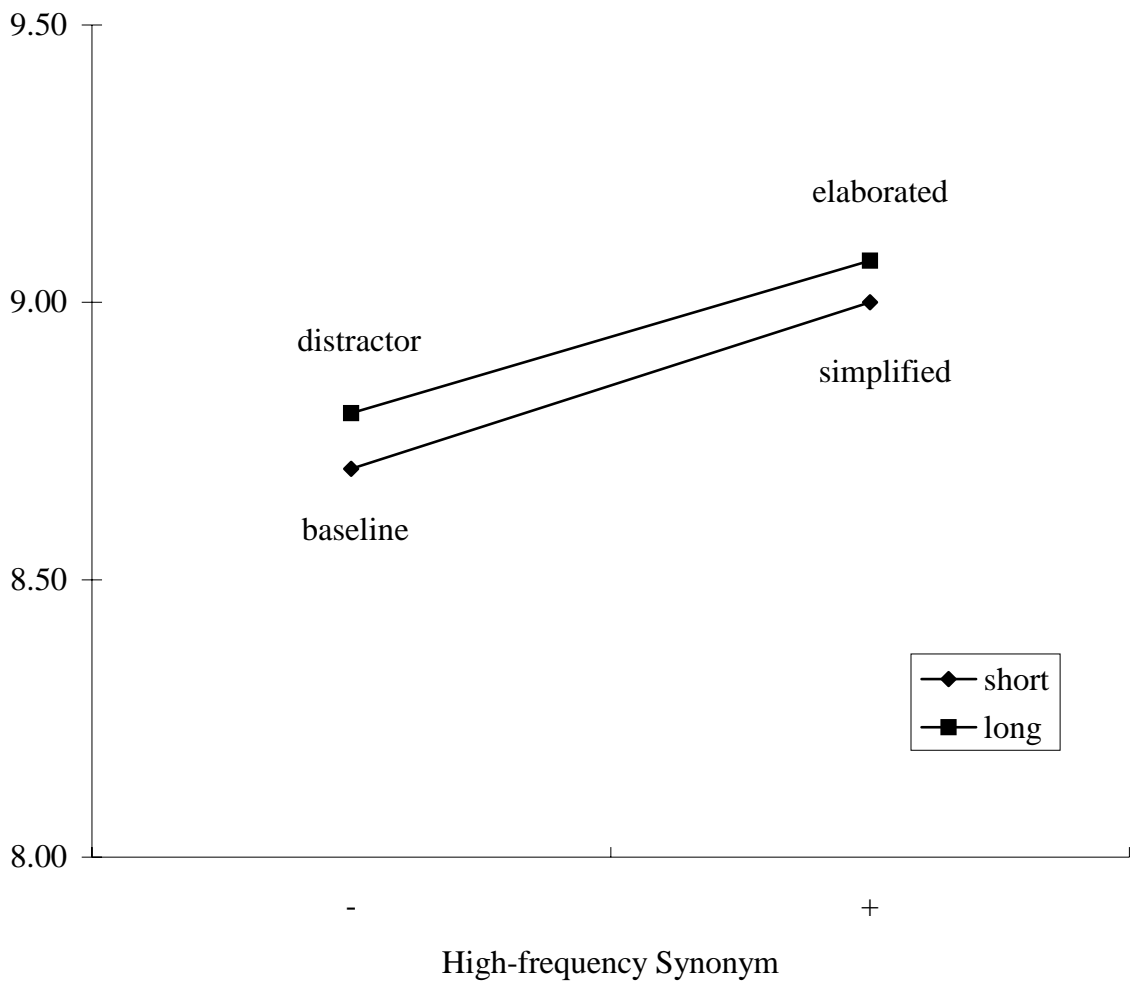


**Figure 3. Mean Reading Times**

**Comprehension questions.** The reliability (Cronbach's alpha) of the overall test was .51. Since the reliability was found to be relatively low, the results should be interpreted with caution. Mean scores and standard deviations for the comprehension questions are presented in Table 4, and the relationships between the two IVs are presented graphically in Figure 4. As Table 4 shows, the mean scores of the four conditions were close to one another. Moreover, they were all found to be high ( $M_s = 8.70 - 9.08$  where  $k = 10$ ), suggesting that their distributions were all negatively skewed. The low reliability of the test is probably due to the skewed distributions (see Brown, 1996, p. 209).

**Table 4**  
*Means and Standard Deviations for Comprehension Scores*  
( $n = 40$ )

Condition	$k$	$M$	$SD$
Baseline	10	8.70	1.02
Elaborated	10	9.08	0.83
Simplified	10	9.00	0.90
Distractor	10	8.80	1.07



**Figure 4. Mean Comprehension Scores**

A multivariate analysis of variance (MANOVA) was performed on the data presented above. All analyses (Pillai's Trace, Wilk's Lambda, Hotelling's Trace, and Roy's Largest Root) indicated multivariate significance for sentence length and modifications and their interaction (see Table 5). Therefore, a univariate analysis of variance (ANOVA) was computed for each dependent variable.

**Table 5***MANOVA Table for Comprehension Measures*

		Value	Approx.	H-df	E-df	p
		<i>F</i>				
Sentence Length	Pillai's Trace	0.75	55.57	2.00	38.00	.000*
	Wilk's Lambda	0.26	55.57	2.00	38.00	.000*
	Hotelling's Trace	2.92	55.57	2.00	38.00	.000*
	Roy's Largest Root	2.92	55.57	2.00	38.00	.000*
Modification	Pillai's Trace	0.61	29.55	2.00	38.00	.000*
	Wilk's Lambda	0.39	29.55	2.00	38.00	.000*
	Hotelling's Trace	1.56	29.55	2.00	38.00	.000*
	Roy's Largest Root	1.56	29.55	2.00	38.00	.000*
Length x Modif.	Pillai's Trace	0.29	7.84	2.00	38.00	.001*
	Wilk's Lambda	0.71	7.84	2.00	38.00	.001*
	Hotelling's Trace	0.41	7.84	2.00	38.00	.001*
	Roy's Largest Root	0.41	7.84	2.00	38.00	.001*

*Notes.* H-df: Hypothesis df, E-df: Error df

\*  $p < .01$

Table 6 presents the results of an ANOVA for mean reading times. The results of the ANOVA indicate significant main effects for both independent variables, i.e., sentence length and lexical modification in the form of the provision of high-frequency synonyms ( $F_s = 101.26$  and  $54.30$ ,  $p_s = .000$ , respectively). A significant interaction was also found ( $F = 15.85$ ,  $p = .000$ ). The strength of association,  $\eta^2$  ( $\eta^2$ ), for the main effects and their interaction was also computed, and found to be .088, .064, and .012, respectively. Despite the significant main effects and interaction, the values  $\eta^2$  were noticeably small. This is partly due to the fact that, in the two-factor repeated-measures design, the total sources of variability were very large, when a measure of association tends to be negatively biased (Girden, 1992, p. 41).

**Table 6*****Two-way Repeated-measures ANOVA Table for Mean Reading Times***

Source	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p</i>	$\eta^2$
Between subjects effects						
Within groups	160313	39	4111			
Within subjects effects						
Sentence length	19445	1	19445	101.26	.000*	.088
Within groups	7489	39	192			
Modification	14178	1	14178	54.30	.000*	.064
Within groups	10183	39	261			
Length x Modification	2588	1	2588	15.85	.000*	.012
Within groups	6366	39	163			
Total	220562	159				

\*  $p < .01$

An ANOVA was conducted for the comprehension scores (Table 7). The results show that sentence length did not affect comprehension scores ( $F = 0.40, p = .531$ ). With regard to the effects of lexical modification, provision of high-frequency synonyms seems to have had a positive influence on sentence comprehension, but the result was not statistically significant ( $F = 4.31, p = .045$ ). No significant interaction was found in this analysis ( $F = 0.01, p = .926$ ).  $Eta^2$  for the main effects and their interaction were found to be minimal.

**Table 7***Two-way Repeated-measures ANOVA Table for Comprehension Scores*

Source	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p</i>	$\eta^2$
Between subjects effects						
Within groups	55.44	39	1.42			
Within subjects effects						
Sentence length	0.31	1	0.31	0.40	.531	.002
Within groups	29.94	39	0.77			
Modification	3.31	1	3.31	4.31	.045	.022
Within groups	29.94	39	0.77			
Length x Modification	0.01	1	0.01	0.01	.926	.000
Within groups	28.24	39	0.72			
Total	147.19	159				

*Vocabulary acquisition measures*

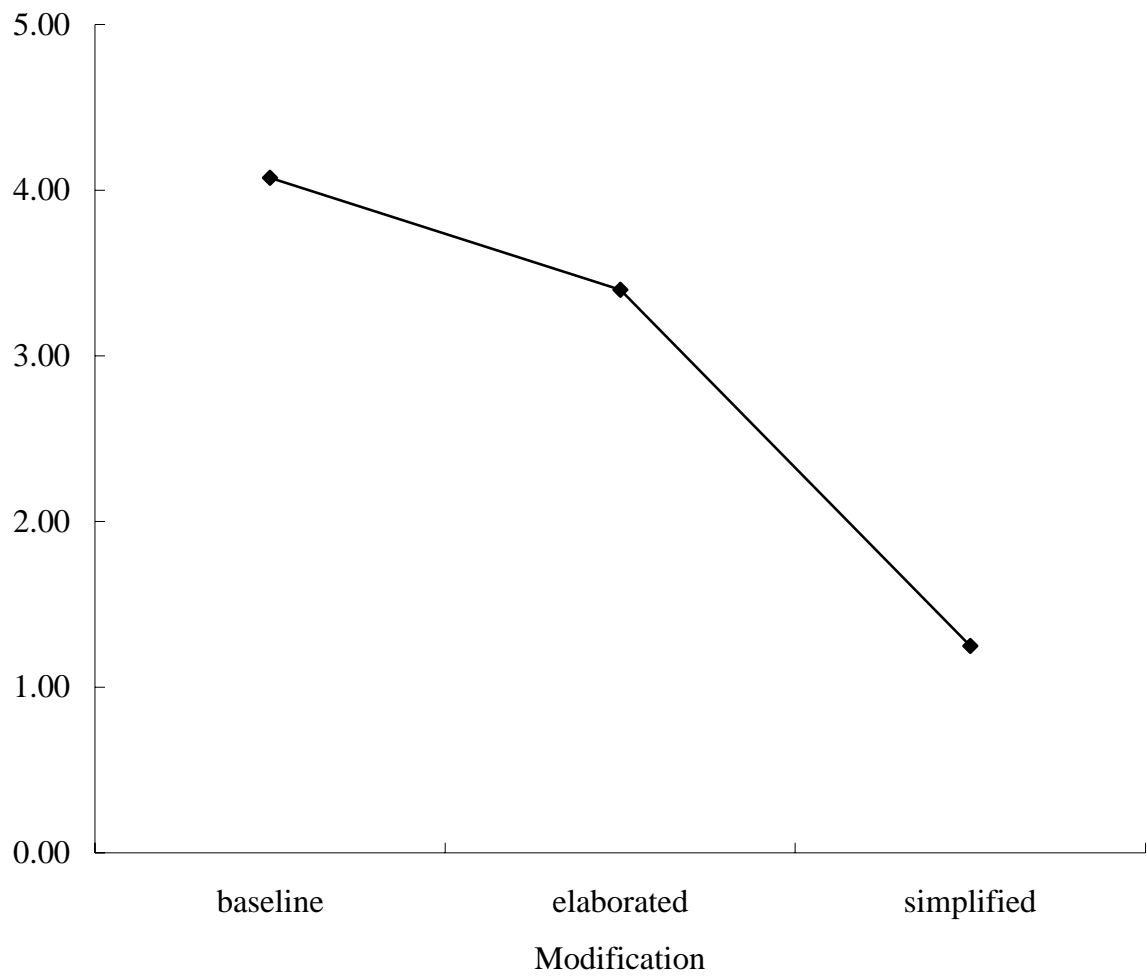
*Form-recognition test.* The overall reliability of the form-recognition test was found to be fairly high ( $\alpha = .80$ ). Table 8 indicates the mean scores and standard deviations for the form-recognition test (see also Figure 5). As shown in Table 8, the mean score for the simplified condition was lower than those for the baseline and elaborated conditions.

**Table 8**

*Form-recognition Test Means and Standard Deviations*

**( $n = 40$ )**

Condition	<i>k</i>	<i>M</i>	<i>SD</i>
Baseline	10	4.08	2.21
Elaborated	10	3.40	2.10
Simplified	10	1.25	1.55

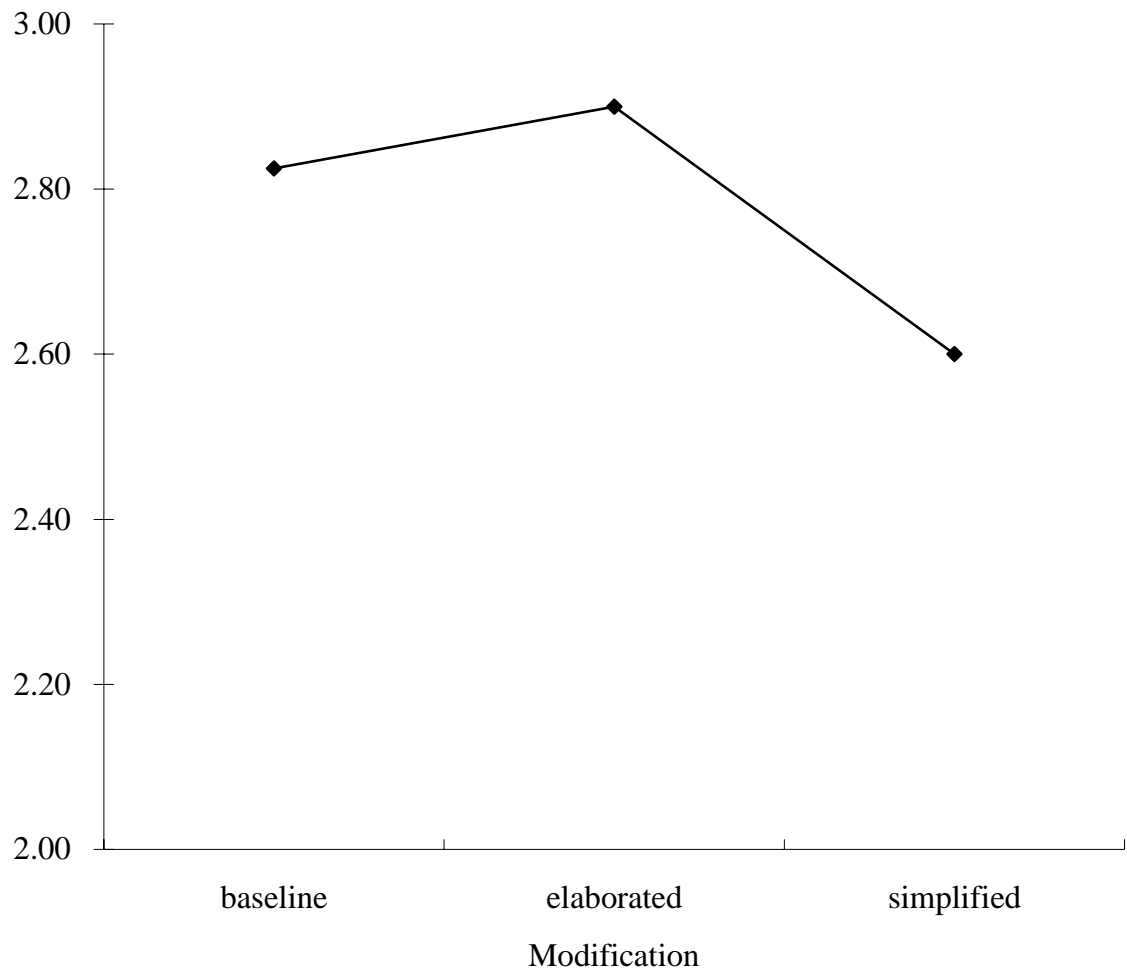


***Figure 5. Form-recognition Test Mean Scores***

**Meaning-recognition test.** Table 9 indicates the mean scores and standard deviations for the meaning-recognition test (See also Figure 6). The differences in scores among the three conditions are smaller than those of the form-recognition test. Also, the reliability of the test, i.e.,  $\alpha = .39$ , was lower than that of the form-recognition test. This low reliability is probably due to the fact that the test was difficult and therefore the distributions were all positively skewed. (See Chapter 5 for further discussion.)

**Table 9**  
***Meaning-recognition Test Means and Standard Deviations***  
**( $n = 40$ )**

Condition	$k$	$M$	$SD$
Baseline	10	2.83	1.43
Elaborated	10	2.90	1.60
Simplified	10	2.60	1.24



***Figure 6. Meaning-recognition Test Mean Scores***

The result of a MANOVA for the vocabulary tests is displayed in Table 1. All analyses (Pillai's Trace, Wilk's Lambda, Hotelling's Trace, and Roy's Largest Root) indicated multivariate significance for the independent variable. Therefore, two one-way repeated-measures ANOVAs were conducted for the form-recognition test and the meaning-recognition test.

**Table 10**  
*MANOVA Table for Vocabulary Tests*

		Value	Approx. <i>F</i>	<i>H-df</i>	<i>E-df</i>	<i>p</i>
Modification	Pillai's Trace	0.48	12.20	4.00	156.00	.000*
	Wilk's Lambda	0.53	14.66	4.00	154.00	.000*
	Hotelling's Trace	0.90	17.17	4.00	152.00	.000*
	Roy's Largest Root	0.90	35.14	2.00	78.00	.000*

*Notes.* *H-df*: Hypothesis df, *E-df*: Error df

\*  $p < .01$

**Table 11*****One-way Repeated-measures ANOVA Table for the Form-recognition Test***

Source	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p</i>	$\eta^2$
Condition (A)	174.12	2	87.06	35.02	.000*	.276
Subjects (S)	261.99	39	6.72			
A x S	193.88	78	2.49			
Total	629.99	119				

\*  $p < .01$

As Table 11 shows, a significant main effect was found for the form-recognition test ( $F = 35.02, p = .000$ ). Therefore, post hoc multiple comparisons were computed to identify the significant mean differences among the three conditions (Table 12).

**Table 12*****Post Hoc Comparisons Table (Paired t-tests using Bonferroni Method)***

	<i>M</i>	<i>SD</i>	<i>t</i>	<i>df</i>	<i>p</i>
Baseline - Elaborated	0.68	2.22	1.92	39	.062
Baseline -Simplified	2.83	2.33	7.67	39	.000*
Elaborated - Simplified	2.15	2.13	6.38	39	.000*

\*  $p < .003 (= .01/3)$

The results of the post hoc comparisons suggest that scores of both the baseline and the

elaborated conditions were significantly higher than that of the simplified conditions. No significant difference was found between the baseline and the elaborated conditions.

Table 13 presents the result of an ANOVA for the meaning-recognition test. Although the scores for the elaborated condition was higher than those for the other conditions, and those of the simplified condition was the lowest, as Hypotheses 8 - 10 assumed, none of the differences was statistically significant ( $F = 0.50, p = .608$ ).

**Table 13**

*One-way Repeated-measures ANOVA Table for Meaning-recognition Test*

Source	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p</i>	$\eta^2$
Condition (A)	1.95	2	0.98	0.50	.608	.008
Subjects (S)	86.93	39	2.23			
A x S	152.05	78	1.95			
Total	240.93	119				

## CHAPTER 5

### DISCUSSION

#### *Hypotheses*

Hypothesis 1 was supported. The mean reading times for the baseline and the simplified conditions were significantly shorter than the other two conditions, i.e., the distractor and the elaborated conditions, as Table 6 shows ( $F = 101.26, p = .000$ ). This suggests that there is some positive relationship between sentence length and reading time.

Hypothesis 2 was also supported. The mean reading times for the modified conditions, i.e., the simplified condition and the elaborated condition, were significantly shorter than those of the other conditions, i.e., the baseline condition and the distractor condition ( $F = 54.30, p = .000$ ). This supports the assumption that lexical modification in the form of providing high-frequency synonyms enhances learner comprehension, thus reducing reading time.

In the analysis of the mean reading times, a significant interaction effect between sentence length and modification was also found ( $F = 15.85, p = .000$ ). This will be discussed later in this chapter.

Hypothesis 3 was supported. The two-way repeated-measures ANOVA for the comprehension scores (Table 7) indicates that the main effect for sentence length was not significant ( $F = 0.40, p = .531$ ). The two-word difference in sentence length did not influence comprehension of the whole sentence.

Hypothesis 4 was not supported. The main effect for modification did not reach statistical significance at the  $p < .01$  level ( $F = 4.31, p = .045$ ). Although the scores of the

modified conditions seem to be consistently higher than those of the other conditions, as Table 4 and Figure 4 show, the difference was not statistically significant.

Hypotheses 5 - 7 were all supported. The mean scores for the form-recognition test in the baseline condition and the elaborated condition were both significantly higher than that in the simplified condition, as the post hoc comparisons demonstrated (Table 12). The difference of 0.68 between the baseline condition and the elaborated condition was not statistically significant, thus supporting Hypothesis 7.

Hypothesis 8 was not supported. Although the mean score for the meaning-recognition test in the elaborated condition was higher than that in the baseline condition by 0.07, this difference was not statistically significant.

Neither was Hypothesis 9 supported. There was a 0.30 difference in the mean scores between the elaborated condition and the simplified condition, as hypothesized, but the difference was not statistically significant.

As stated in Hypothesis 10, there was no significant difference between the baseline condition and the simplified condition in the mean scores for the meaning-recognition test. However, this may not be important since no clear tendency was found in scores on the meaning-recognition test as a whole.

### ***Lexical simplification, elaboration, and sentence comprehension***

Results of this study have demonstrated positive effects of lexical simplification and elaboration on L2 comprehension at the sentential level. Providing high-frequency words either in place of or in apposition to the target unknown words significantly shortened mean reading times, as the result of an ANOVA shows (Table 6). This confirms general

tendencies for results of text modification in previous studies: both simplification and elaboration improve L2 comprehension.

Table 6 also shows that there is a significant interaction between sentence length and modification, which needs interpretation. As Table 3 and Figure 3 indicate, the difference in the mean reading times between the baseline and the simplified conditions (i.e., 26.87 seconds) is larger than the difference between the distractor and the elaborated conditions (10.78 seconds). One potential interpretation of this difference is that lexical simplification has a stronger impact on comprehension than lexical elaboration does, thus shortened the mean reading time more substantially. This interpretation does not conflict with previous findings. Three of the four studies reviewed in Chapter 2 (Table 1) found stronger effects for simplification than elaboration, although the differences were not always statistically significant. It is possible that simplification has a stronger impact on sentence comprehension since it involves mostly decoding of the literal meaning. Several studies have found this tendency (e.g., Chaudron, 1983; Long, 1985).

There is another plausible interpretation, however. The most remarkable difference between the simplified and the elaborated conditions is that the target lexical items were removed from the simplified sentences, whereas they were retained in the elaborated sentences. Reading elaborated sentences, therefore, requires another psychological process of recognizing the presence of unknown words in the sentences. Put another way, it can be assumed that participants did more than mere reading of sentences for comprehension when they saw the elaborated sentences. Some mental process of form-meaning matching of the unknown words might have taken place while they read the elaborated sentences, and thus led to the relatively smaller effect of lexical elaboration than

simplification. All of these interpretations, however, are no more than speculations without supporting evidence, and further research would be necessary to confirm or reject them.

The comprehension scores have provided additional supports for the positive effects of both simplification and elaboration on comprehension. Although it did not reach the statistical significance at the  $p < .01$  level, an ANOVA showed a tendency for a positive effect of lexical modification on comprehension ( $F = 4.31, p = .045$ ). There was little difference in scores between the simplified condition and the elaborated condition, and both were higher than the scores in the baseline and the distractor conditions (See Table 4 & Figure 4). It is important to note that the comprehension questions were formulated so that they could be answered correctly without knowing the meaning of the target words. In other words, whether or not high-frequency synonyms were provided did not influence comprehension scores since the questions were targeted at other parts of the test sentences, not the unknown words themselves. The lower scores in the baseline and distractor conditions may suggest that encountering an unknown word without any clue as to its meaning distracts the learner's attention from other parts of the sentence. This possibility also awaits further research.

### ***Lexical simplification, elaboration, and incidental vocabulary acquisition***

The results of the form-recognition test were straightforward. The mean scores in the baseline condition and the elaborated condition were significantly higher than that in the simplified condition (See Table 12 for the results of the post hoc pairwise comparisons). This suggests that presentation of target lexical items itself can trigger the first step of

acquisition, i.e., recognition of form. This supports Long's argument in favor of elaboration over simplification (Long, 1983; 1996; Yano et al., 1994). If both simplification and elaboration enhance comprehension, which was the case in the present study, then elaboration is a preferable type of modification since it provides L2 learners with opportunities to learn new linguistic items.

On the other hand, results of the meaning-recognition test did not show the same tendency clearly. The mean score in the elaborated condition was slightly higher than the other two, and that in the simplified condition was the lowest, as Hypotheses 8 - 10 predicted; however, the differences were not statistically significant (See Table 13). The mean scores in the baseline, elaborated, and distractor conditions were 2.83, 2.90, and 2.60, respectively. Since the test consisted of one-out-of-four multiple-choice questions, an average of 2.50 could be achieved by chance alone on ten test items ( $10 \times 1/4 = 2.5$ ). Therefore, these mean scores are not much higher than chance. The slightly higher score in the elaborated condition might have been the result of lexical elaboration promoting vocabulary acquisition, but the effect found in the experiment was minimal. A possible explanation for this null finding is that one occurrence of each target word was insufficient for this level of vocabulary learning to take place. As mentioned earlier, recognition of word form is considered to come earlier than recognition of word meaning, and therefore exposure to more than one occurrence may be necessary for successful acquisition. Results of previous studies of incidental vocabulary acquisition are varied, but at least several occurrences seem to be necessary. (See Huckin & Coady, 1999, for a recent review of this issue.)

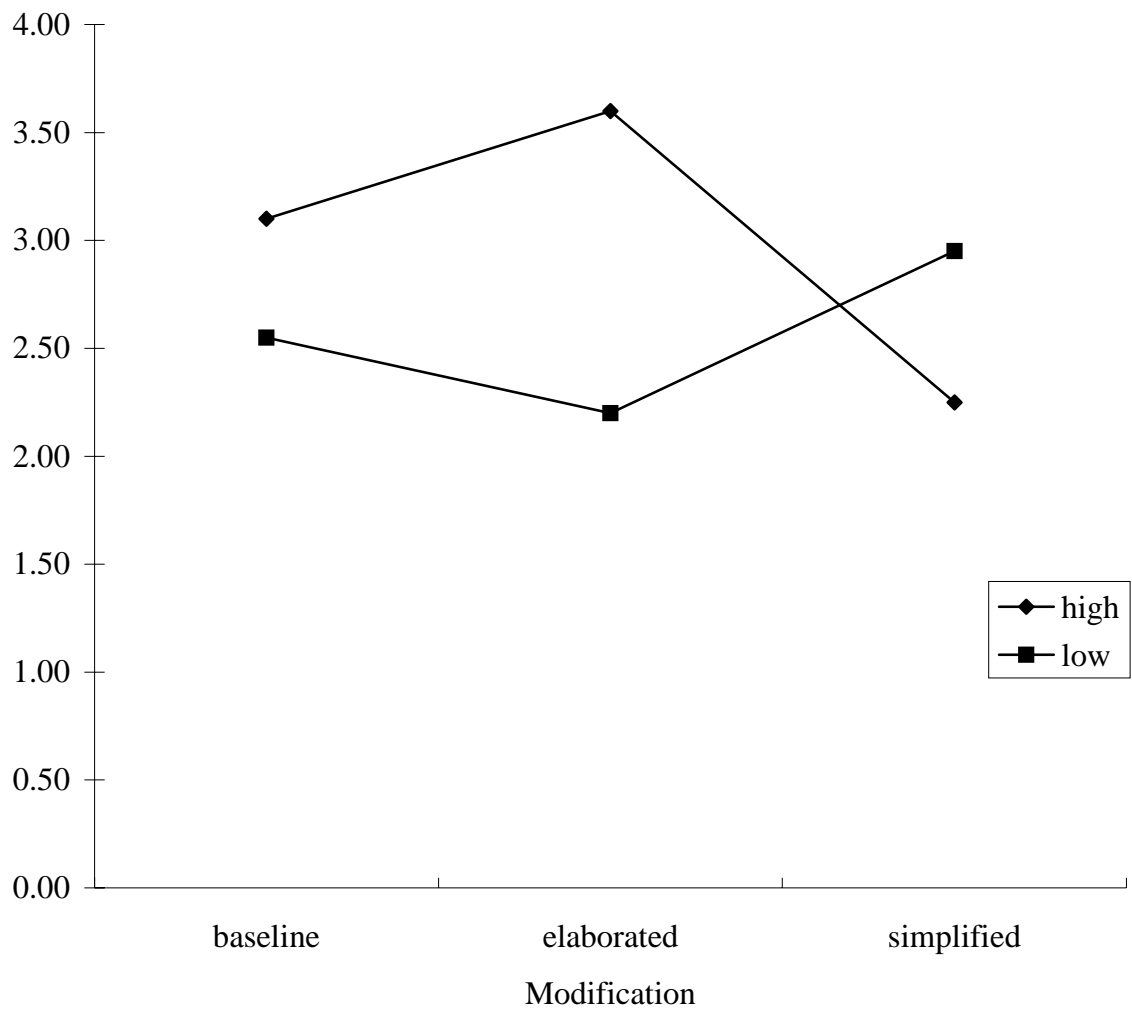
There is another possibility that might account for the result. It is possible that learners

of a certain proficiency level benefit more from lexical elaboration than others. Since proficiency in the present study ranged fairly widely (e.g., TOEFL scores between 350 and 573), differences in proficiency might have obscured the results of the meaning-recognition test. Therefore, an additional analysis was conducted on the meaning-recognition test scores to investigate the influence of proficiency. Participants were divided into two groups according to proficiency, using scores on a cloze test administered in the experimental session. The cloze test was created by Brown, Yamashiro, and Ogane (1999), and was tailored for Japanese university students. The reliability of the thirty-item test in the present study was found to be .68, using Cronbach's alpha.

Participants were ranked by their cloze scores, and then divided into two groups, i.e., high- and low-proficiency. Table 14 shows means and standard deviations for the meaning recognition-test scores (see also Figure 7).

**Table 14***Means and Standard Deviations for Meaning-recognition Test Scores**for Two Proficiency Levels**(k = 10)*

Condition	<i>n</i>		<i>M</i>		<i>SD</i>	
	High	Low	High	Low	High	Low
Baseline	20	20	3.10	2.55	1.07	1.70
Elaborated	20	20	3.60	2.20	1.43	1.47
Simplified	20	20	2.25	2.95	1.12	1.28



**Figure 7. Mean Meaning-recognition Test Scores  
for Two Proficiency Levels**

A 2 x 3 repeated-measures ANOVA was conducted with proficiency and lexical modifications as independent variables and meaning-recognition test scores as the dependent variable. The result is presented in Table 15.

**Table 15**  
*Two-way Repeated-measures ANOVA Table for Meaning-recognition Test  
for Two Proficiency Levels*

Source	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p</i>	$\eta^2$
Between subjects effects						
Proficiency	5.21	1	5.21	2.42	.128	.022
Within groups	81.72	38	2.15			
Within subjects effects						
Modification	1.95	2	0.57	0.57	.567	.008
Proficiency x Modif.	22.32	2	11.16	6.54	.002*	.093
Within groups	129.73	76	1.71			
Total	240.93	119				

\*  $p < .01$

The main effect for proficiency was not significant at the  $p < .01$  level. However, it is important to note that the interaction between proficiency and modification was statistically significant ( $F = 6.54, p = .002$ ). This interaction, which is also displayed in Figure 7, suggests that participants of different proficiency levels were influenced

differently by lexical modification. Figure 7 shows that performance of the high-proficiency group was closer to what was hypothesized: highest scores in the elaborated condition, and the lowest in the simplified condition. Post hoc comparisons among the participants of the same proficiency level were also computed (Tables 16 & 17).

**Table 16**

*Post Hoc Comparisons Table (Paired t-tests using Bonferroni Method)*

*High-proficiency Group*

	<i>M</i>	<i>SD</i>	<i>t</i>	<i>df</i>	<i>p</i>
Baseline - Elaborated	- 0.50	1.57	- 1.42	19	.171
Baseline - Simplified	0.85	1.23	3.10	19	.006
Elaborated - Simplified	1.35	2.01	3.07	19	.007

**Table 17**

*Post Hoc Comparisons Table (Paired t-tests using Bonferroni Method)*

*Low-proficiency Group*

	<i>M</i>	<i>SD</i>	<i>t</i>	<i>df</i>	<i>p</i>
Baseline - Elaborated	0.35	2.35	0.67	19	.513
Baseline - Simplified	- 0.40	2.04	- 0.88	19	.391
Elaborated - Simplified	- 0.75	1.68	- 1.94	19	.061

Table 16 shows that, within the high-proficiency group, the mean score in the simplified condition might be lower than scores in both the baseline and elaborated

conditions in a more powerful study, suggesting that lexical simplification deprived participants of opportunities to learn the meanings of target words. The 0.50 difference between the baseline and elaborated conditions may have been due to lexical elaboration enhancing vocabulary acquisition; however, the difference was not statistically significant ( $p = .171$  at the significance level of  $\alpha < .002$  ( $\alpha/6 = .01/6 = .002$ )). The lack of statistical significance may be due to the possibility of incidental vocabulary learning from context in the baseline condition. The present study has not provided enough evidence to assess this possibility, for which further research would be required. No significant difference was found within the low-proficiency group, however, implying that scores in the low-proficiency group were obtained by chance alone.

Results of this additional analysis may suggest that the high-proficiency group benefited more from lexical elaboration, but again the results are not clear enough to be conclusive. It would be legitimate to claim, however, that lexical modification has differential effects on learners of different proficiency.

## CHAPTER 6

### CONCLUSION

#### *Summary of the findings and implications*

The present study investigated the effects of lexical simplification and elaboration on L2 sentence comprehension and incidental vocabulary acquisition. Performance of the forty participants of the study suggests that (a) both lexical simplification and elaboration facilitate L2 comprehension at the sentential level; (b) elaborating target lexical items can trigger acquisition of those elaborated words, whereas lexical simplification hinders acquisition of the target words; and (c) one occurrence of each target word with lexical elaboration may be insufficient for acquisition of its meaning, although more advanced learners may be able to learn some of the words on the basis of only one encounter. These findings are consistent with general tendencies found in previous studies of text modification, L2 comprehension, and vocabulary acquisition.

Although the present study was purely experimental in nature, the findings may have implications for pedagogy. Positive effects of lexical elaboration on both comprehension and vocabulary acquisition found in the present study lend support to the use of this technique in designing L2 reading materials. An advantage of lexical elaboration is that it can be used with relative ease: adding synonyms to supposedly unknown words will not require much time and effort on the part of the language teacher and/or text writer, yet positive effects on both comprehension and vocabulary acquisition can be expected. Another advantage exists for the L2 learner. Although lexical elaboration lengthens sentences, it does not seem to add much of a burden for L2 learners. In the present study,

the elaborated sentences were roughly 15% longer than the baseline sentences (a two-word difference in the mean sentence length of 13.7 words). Nevertheless, the difference in mean reading times was relatively small (3.22 seconds, or approximately 3%, of the mean reading time in the baseline condition), suggesting that improved comprehension and possibly some vocabulary acquisition can be achieved without much extra effort.

### *Limitations of the study and suggestions for future studies*

First, it should be noted that the experiment took place under conditions that were different from natural L2 reading/learning environments. The participants knew they were being tested as part of an experiment, although they were not informed about its exact purpose. Second, the reading was done sentence by sentence on a computer screen, followed by comprehension questions, which is rarely the case in real-life reading situations. There is a possibility that such differences might have influenced the participants in some ways. Third, acquisition of the forms and meanings of nonsense words may be different from learning existing vocabulary items in a target language. Fourth, the specific participant population, native speakers of Japanese learning English in ESL situations, limits the generalizability of the findings. Replication studies with larger and more varied populations will expand the scope of the findings.

Fifth, some of the measures employed in the present study were found to be less reliable than others. For example, the reliability of the meaning-recognition test ( $\alpha = .39$ ) was much lower than that of the other measures. The lack of statistical significance in the meaning-recognition test could have been due to this low reliability. In other words, it is possible that there were other factors that influenced the results, thus obscuring the effect

of the independent variable.

Suggestions for future studies are as follows. First, replication studies are necessary to confirm the findings of the present work. More powerful studies with larger *n*-sizes would provide a clearer picture of the effects of lexical modification on reading comprehension and vocabulary acquisition. Also, since it was not clear from the present study whether lexical elaboration promoted acquisition of word meanings, it is necessary to conduct studies with more than one occurrence of the target vocabulary items. Also, studies of lexical elaboration and vocabulary acquisition in regular instructional settings would provide more pertinent information for pedagogy. Longitudinal classroom research with more realistic reading materials and vocabulary items would expand the applicability of the findings to L2 teaching. In like manner, it would be of some theoretical and pedagogical interest to investigate effects of lexical modification on L2 learners of different proficiency levels.

## NOTES

1. It should be noted that one of the two professors in this study did not show this tendency. Chaudron (1988) states that there is no knowing whether the difference between the two professors is due to (a) variability within the professors observer, (b) different language-specific norms for speech rate between English and French, or (c) the instructional context (p. 69).
2. This includes institutional TOEFL which is administered by ESL programs affiliated with the University of Hawai'i.
3. The temporal resolution of the Macintosh operating system allows a minimum of 16 milliseconds accuracy for the timing events (Cohen et al., 1993, p. 269).

**APPENDIX A**  
**TEST SENTENCES AND COMPREHENSION QUESTIONS**

1. B(aseline): Kyoko asked Jerry the plear of the mountain that was covered with snow.  
S(implified): Kyoko asked Jerry the height of the mountain that was covered with snow.  
E(laborated): Kyoko asked Jerry the plear, or height, of the mountain that was covered with snow.  
D(istractor): Kyoko asked Jerry the plear, or hister, of the mountain that was covered with snow.  
Q(uestion): Did Kyoko see the mountain?
  
2. B: Because of the darm weather, Kevin does not like the rainy season very much.  
S: Because of the wet weather, Kevin does not like the rainy season very much.  
E: Because of the darm, or wet, weather, Kevin does not like the rainy season very much.  
D: Because of the darm, or splim, weather, Kevin does not like the rainy season very much.  
Q: Does Kevin like the rainy season?
  
3. B: Joan bought an ober and a pair of shoes when she went shopping last Sunday.  
S: Joan bought a bag and a pair of shoes when she went shopping last Sunday.  
E: Joan bought an ober, or bag, and a pair of shoes when she went shopping last Sunday.  
D: Joan bought an ober, or hinch, and a pair of shoes when she went shopping last Sunday.  
Q: Did Joan go shopping last Friday?
  
4. B: Lisa could not stop smubbing because she was watching a very sad movie.  
S: Lisa could not stop crying because she was watching a very sad movie.  
E: Lisa could not stop smubbing, or crying, because she was watching a very sad movie.  
D: Lisa could not stop smubbing, or initating, because she was watching a very sad movie.  
Q: Was the movie sad?
  
5. B: Because Linda wants to ungle a new car, she has to work more.  
S: Because Linda wants to buy a new car, she has to work more.  
E: Because Linda wants to ungle, or buy, a new car, she has to work more.  
D: Because Linda wants to ungle, or vonce, a new car, she has to work more.  
Q: Does Linda have to work more?
  
6. B: Mary saw a small rascine singing when she looked outside through the window.  
S: Mary saw a small bird singing when she looked outside through the window.  
E: Mary saw a small rascine, or bird, singing when she looked outside through the window.  
D: Mary saw a small rascine, or ruspence, singing when she looked outside through the window.  
Q: Did Mary go outside?

7. B: Because Robert was very litidate to Rose at the meeting, she likes him.  
 S: Because Robert was very kind to Rose at the meeting, she likes him.  
 E: Because Robert was very litidate, or kind, to Rose at the meeting, she likes him.  
 D: Because Robert was very litidate, or feach, to Rose at the meeting, she likes him.  
 Q: Does Rose like Robert?
8. B: Because it did not rimmor raining yesterday, Vivian could not go out all day.  
 S: Because it did not stop raining yesterday, Vivian could not go out all day.  
 E: Because it did not rimmor, or stop, raining yesterday, Vivian could not go out all day.  
 D: Because it did not rimmor, or tiver, raining yesterday, Vivian could not go out all day.  
 Q: Did Vivian go out yesterday?
9. B: Craig did not praverse someone stealing his money because the room was very dark.  
 S: Craig did not notice someone stealing his money because the room was very dark.  
 E: Craig did not praverse, or notice, someone stealing his money because the room was very dark.  
 D: Craig did not praverse, or torth, someone stealing his money because the room was very dark.  
 Q: Was the room very dark?
10. B: Because his pencil was not moast, Mike asked Tony to lend him one.  
 S: Because his pencil was not sharp, Mike asked Tony to lend him one.  
 E: Because his pencil was not moast, or sharp, Mike asked Tony to lend him one.  
 D: Because his pencil was not moast, or armot, Mike asked Tony to lend him one.  
 Q: Did Mike ask Tony for a pencil?
11. B: Cathy was reading a magazine when she saw a serald right in front of her.  
 S: Cathy was reading a magazine when she saw an accident right in front of her.  
 E: Cathy was reading a magazine when she saw a serald, or accident, right in front of her.  
 D: Cathy was reading a magazine when she saw a serald, or cratic, right in front of her.  
 Q: Was Cathy reading a newspaper?
12. B: Ken did not know Kate was in the kitchen because she was very marfle.  
 S: Ken did not know Kate was in the kitchen because she was very quiet.  
 E: Ken did not know Kate was in the kitchen because she was very marfle, or quiet.  
 D: Ken did not know Kate was in the kitchen because she was very marfle, or golt.  
 Q: Was Kate in the kitchen?
13. B: This morning Kim went to the supermarket to buy some butty and milk.  
 S: This morning Kim went to the supermarket to buy some bread and milk.  
 E: This morning Kim went to the supermarket to buy some butty, or bread, and milk.  
 D: This morning Kim went to the supermarket to buy some butty, or lipal, and milk.  
 Q: Did Kim go to the bookstore this morning?
14. B: Kevin does not like his new apartment because he thinks it is very exlume.  
 S: Kevin does not like his new apartment because he thinks it is very dirty.  
 E: Kevin does not like his new apartment because he thinks it is very exlume, or dirty.  
 D: Kevin does not like his new apartment because he thinks it is very exlume, or tostle.  
 Q: Does Kevin like his new apartment?

15. B: Jennifer saw Takashi and his friends at the nace to the hotel last night.  
 S: Jennifer saw Takashi and his friends at the entrance to the hotel last night.  
 E: Jennifer saw Takashi and his friends at the nace, or entrance, to the hotel last night.  
 D: Jennifer saw Takashi and his friends at the nace, or soor, to the hotel last night.  
 Q: Did Jennifer see Takashi last night?
16. B: Robert was very scared of the animal because it was very big and selest.  
 S: Robert was very scared of the animal because it was very big and hairy.  
 E: Robert was very scared of the animal because it was very big and selest, or hairy.  
 D: Robert was very scared of the animal because it was very big and selest, or telt.  
 Q: Was Robert scared of the animal?
17. B: Kevin is very dield of Ken because he is rich, smart, and handsome.  
 S: Kevin is very jealous of Ken because he is rich, smart, and handsome.  
 E: Kevin is very dield, or jealous, of Ken because he is rich, smart, and handsome.  
 D: Kevin is very dield, or nooch, of Ken because he is rich, smart, and handsome.  
 Q: Is Ken handsome?
18. B: Craig could not sleep well last night because of the gank outside his house.  
 S: Craig could not sleep well last night because of the noise outside his house.  
 E: Craig could not sleep well last night because of the gank, or noise, outside his house.  
 D: Craig could not sleep well last night because of the gank, or shapper, outside his house.  
 Q: Did Craig sleep well last night?
19. B: Jerry likes his house because he can see the cermit from the living room.  
 S: Jerry likes his house because he can see the ocean from the living room.  
 E: Jerry likes his house because he can see the cermit, or ocean, from the living room.  
 D: Jerry likes his house because he can see the cermit, or finter, from the living room.  
 Q: Does Jerry like his house?
20. B: Tom tepsined me he would come to the meeting, but he did not show up.  
 S: Tom promised me he would come to the meeting, but he did not show up.  
 E: Tom tepsined, or promised, me he would come to the meeting, but he did not show up.  
 D: Tom tepsined, or griled, me he would come to the meeting, but he did not show up.  
 Q: Did Tom come to the meeting?
21. B: Kim is not happy because there are very few good sleams near her house.  
 S: Kim is not happy because there are very few good restaurants near her house.  
 E: Kim is not happy because there are very few good sleams, or restaurants, near her house.  
 D: Kim is not happy because there are very few good sleams, or brudges, near her house.  
 Q: Is Kim happy?
22. B: Liz didn't want to talk to Bill because he was very rigure to her yesterday.  
 S: Liz didn't want to talk to Bill because he was very rude to her yesterday.  
 E: Liz didn't want to talk to Bill because he was very rigure, or rude, to her yesterday.  
 D: Liz didn't want to talk to Bill because he was very rigure, or smour, to her yesterday.  
 Q: Did Liz want to talk to Bill?

23. B: Andrew cannot titch Diana any more because she has lied to him before.  
 S: Andrew cannot trust Diana any more because she has lied to him before.  
 E: Andrew cannot titch, or trust, Diana any more because she has lied to him before.  
 D: Andrew cannot titch, or rastle, Diana any more because she has lied to him before.  
 Q: Has Diana told a lie to Andrew?
24. B: Alan wants to move to the south because the weather is smaper there.  
 S: Alan wants to move to the south because the weather is warmer there.  
 E: Alan wants to move to the south because the weather is smaper, or warmer, there.  
 D: Alan wants to move to the south because the weather is smaper, or soaber, there.  
 Q: Does Alan want to move to the south?
25. B: Linda couldn't go to the party yesterday because her parents didn't elert her to.  
 S: Linda couldn't go to the party yesterday because her parents didn't allow her to.  
 E: Linda couldn't go to the party yesterday because her parents didn't elert, or allow, her to.  
 D: Linda couldn't go to the party yesterday because her parents didn't elert, or treize, her to.  
 Q: Did Linda go to the party yesterday?
26. B: Richard conpealed that his office does not have a connection to the internet.  
 S: Richard complained that his office does not have a connection to the internet.  
 E: Richard conpealed, or complained, that his office does not have a connection to the internet.  
 D: Richard conpealed, or slired, that his office does not have a connection to the internet.  
 Q: Does Richard's office have a connection to the internet?
27. B: Brian works for a big mostage which is near Ala Moana Center every weekday.  
 S: Brian works for a big factory which is near Ala Moana Center every weekday.  
 E: Brian works for a big mostage, or factory, which is near Ala Moana Center every weekday.  
 D: Brian works for a big mostage, or anept, which is near Ala Moana Center every weekday.  
 Q: Does Brian work on weekends?
28. B: John wants to build a new pardel because he recently bought a new car.  
 S: John wants to build a new garage because he recently bought a new car.  
 E: John wants to build a new pardel, or garage, because he recently bought a new car.  
 D: John wants to build a new pardel, or anvoy, because he recently bought a new car.  
 Q: Did John buy a new car recently?
29. B: Charlie didn't change his mind at all even though Nick tried to raste him to.  
 S: Charlie didn't change his mind at all even though Nick tried to persuade him to.  
 E: Charlie didn't change his mind at all even though Nick tried to raste, or persuade, him to.  
 D: Charlie didn't change his mind at all even though Nick tried to raste, or crail, him to.  
 Q: Did Charlie change his mind?

30. B: Steve is ausiting something special since his friends will visit him this evening.  
 S: Steve is cooking something special since his friends will visit him this evening.  
 E: Steve is ausiting, or cooking, something special since his friends will visit him this evening.  
 D: Steve is ausiting, or prulicing, something special since his friends will visit him this evening.  
 Q: Will Steve's friends visit him this evening?
31. B: Even though Peter did not agree, Martin thought the movie was very pollow.  
 S: Even though Peter did not agree, Martin thought the movie was very interesting.  
 E: Even though Peter did not agree, Martin thought the movie was very pollow, or interesting.  
 D: Even though Peter did not agree, Martin thought the movie was very pollow, or fick.  
 Q: Did Peter agree with Martin's opinion?
32. B: Jonathan could not stop lassaging when he saw the comedian fall off the stage.  
 S: Jonathan could not stop laughing when he saw the comedian fall off the stage.  
 E: Jonathan could not stop lassaging, or laughing, when he saw the comedian fall off the stage.  
 D: Jonathan could not stop lassaging, or sporling, when he saw the comedian fall off the stage.  
 Q: Did the comedian fall off the stage?
33. B: The doctor told Daniel to take doral every day to cure his illness.  
 S: The doctor told Daniel to take medicine every day to cure his illness.  
 E: The doctor told Daniel to take doral, or medicine, every day to cure his illness.  
 D: The doctor told Daniel to take doral, or midate, every day to cure his illness.  
 Q: Is Daniel sick?
34. B: On her way home off Waikiki, Debbie saw a new dail in the grass.  
 S: On her way home off Waikiki, Debbie saw a new hat in the grass.  
 E: On her way home off Waikiki, Debbie saw a new dail, or hat, in the grass.  
 D: On her way home off Waikiki, Debbie saw a new dail, or benile, in the grass.  
 Q: Does Debbie live in Waikiki?
35. B: Peter could not tundle the cat because it moved very fast and ran away.  
 S: Peter could not catch the cat because it moved very fast and ran away.  
 E: Peter could not tundle, or catch, the cat because it moved very fast and ran away.  
 D: Peter could not tundle, or birb, the cat because it moved very fast and ran away.  
 Q: Did the cat move fast?
36. B: Keith likes the talk show on TV because the comedian is very incrint.  
 S: Keith likes the talk show on TV because the comedian is very funny.  
 E: Keith likes the talk show on TV because the comedian is very incrint, or funny.  
 D: Keith likes the talk show on TV because the comedian is very incrint, or alect.  
 Q: Does Keith like the talk show?

37. B: Keiko studies English very hard because she wants to start reading English rebairs.  
S: Keiko studies English very hard because she wants to start reading English poems.  
E: Keiko studies English very hard because she wants to start reading English rebairs, or poems.  
D: Keiko studies English very hard because she wants to start reading English rebairs, or megates.  
Q: Does Keiko study English hard?
38. B: Yuka decided not to go to the store again because their progue was so bad.  
S: Yuka decided not to go to the store again because their service was so bad.  
E: Yuka decided not to go to the store again because their progue, or service, was so bad.  
D: Yuka decided not to go to the store again because their progue, or ragget, was so bad.  
Q: Would Yuka go to the store again?
39. B: Meg's parents fermed her new house yesterday because it was her 25th birthday.  
S: Meg's parents visited her new house yesterday because it was her 25th birthday.  
E: Meg's parents fermed, or visited, her new house yesterday because it was her 25th birthday.  
D: Meg's parents fermed, or besterned, her new house yesterday because it was her 25th birthday.  
Q: Was yesterday Meg's birthday?
40. B: Becky could not eat the cake because it was too vight for her.  
S: Becky could not eat the cake because it was too sweet for her.  
E: Becky could not eat the cake because it was too vight, or sweet, for her.  
D: Becky could not eat the cake because it was too vight, or malash, for her.  
Q: Did Becky eat the cake?

**APPENDIX B**  
**FORM-RECOGNITION TEST**

下の単語のうち、あなたが先ほどコンピュータの画面で見たと思うものすべてを丸( )で囲んでください。

<例>	window	←見た	door	←見てない
sleam(s)	rascine		pollow	butty
rebair(s)	darm		dield	conpeal(ed)
praverse	vight		raste	titch
nace	ungle		smape(r)	plear
lassage [lassaging]	rigure		elert	exlume
doral	serald		mostage	smub(bing)
rimmor	incrint		moast	cermit
litidate	tundle		ausit(ing)	progue
pardel	marfle		ferm(ed)	ober
selest	tepsin(ed)		dail	gank

自分の見た単語すべてに をつけたと思ったら、次のページに進んでください。

**APPENDIX C**  
**MEANING-RECOGNITION TEST**

下の単語の日本語訳としてもっともふさわしいと思うものをそれぞれ選択肢 (1) - (4) から選んで、番号を で囲んでください。自信のないものも推測して、すべて解答してください。

<例>		desk	(1) 食べる (2) 机 (3) 寝る (4) 怒る		
plear	(1) 騒音 (2) 笑う (3) 高さ (4) 暖かい	ungle	(1) 親切的な (2) 買う (3) 嫉妬する (4) 針	praverse	(1) 気づく (2) さまよう (3) 説得する (4) カバン
butty	(1) 不平を言う (2) 訪れる (3) パン (4) 海	dield	(1) 庭 (2) 量 (3) 料理する (4) 嫉妬する	sleam	(1) 湿った (2) 岸 (3) レストラン (4) 習慣
elert	(1) 甘い (2) 面白い (3) 許可する (4) 鳥	raste	(1) 約束する (2) 距離 (3) 鋭い (4) 説得する	doral	(1) 心配する (2) 帽子 (3) 薬 (4) 事故
rebar	(1) 逃げる (2) 詩 (3) 毛むくじゃらな (4) 泣く	darm	(1) 湿った (2) 静かな (3) 工場 (4) 獲得する	rascine	(1) 糸 (2) 信用する (3) 鳥 (4) ばかな
moast	(1) 止まる (2) 鋭い (3) つかまえる (4) 靴(くつ)	exlume	(1) 汚い (2) 勇敢な (3) 増える (4) 安全な	gank	(1) 騒音 (2) 完璧な (3) 帽子 (4) 薬
rigure	(1) 話す (2) 失礼な (3) 気づく (4) 入口	conpeal	(1) 量 (2) 車庫 (3) 取り除く (4) 不平を言う	ausit	(1) サービス (2) おかしな (3) 料理する (4) 高さ

次のページに続きます。

dail	(1) 詩 (2) 帽子 (3) 海 (4) 庭	progue	(1) 笑う (2) 訪れる (3) サービス (4) 事故	ober	(1) カバン (2) 静かな (3) 獲得する (4) 買う
litidate	(1) さまよう (2) 糸 (3) 勇敢な (4) 親切的な	serald	(1) 事故 (2) 面白い (3) 習慣 (4) 距離	nace	(1) 安全な (2) 取り除く (3) 入口 (4) ばかな
cermit	(1) 話す (2) 許可する (3) 心配する (4) 海	titch	(1) 集める (2) パン (3) 信用する (4) おかしな	mostage	(1) 岸 (2) 工場 (3) 甘い (4) 糸
pollow	(1) 嫉妬する (2) 逃げる (3) 薬 (4) 面白い	tundle	(1) つかまえる (2) 量 (3) 針 (4) 暖かい	ferm	(1) 庭 (2) 車庫 (3) レストラン (4) 訪れる
smub	(1) 泣く (2) 靴(くつ) (3) 増える (4) 約束する	rimmor	(1) 失礼な (2) 汚い (3) 止まる (4) 完璧な	marfle	(1) 習慣 (2) 集める (3) 静かな (4) 買う
selest	(1) 毛むくじゃらな (2) 距離 (3) カバン (4) 逃げる	tepsin	(1) 鋭い (2) 安全な (3) さまよう (4) 約束する	smape	(1) 暖かい (2) 鳥 (3) 取り除く (4) 話す
pardel	(1) 詩 (2) 料理する (3) 不平を言う (4) 車庫	lassage	(1) 泣く (2) 針 (3) 獲得する (4) 笑う	incrint	(1) 工場 (2) おかしな (3) 止まる (4) 岸
vight	(1) ばかな (2) 高さ (3) 甘い (4) 湿った				

## APPENDIX D INSTRUCTIONS FOR THE EXPERIMENT

Aloha! and welcome to my experiment. In this experiment, you will read English sentences and answer simple yes-no questions.

For example, you will read a sentence like:

"Ken is sad because he has lost his money."

And a yes-no question like:

"Is Ken happy?"

If your answer is YES, please hit "y" on the keyboard.

If your answer is NO, please hit "n".

In this example, the answer is no (right?), so you need to hit "n" key. OK?

If you have any question so far, please raise your hand and ask me.

If not, please hit "x" key to move on.

When you start the experiment, you will find a sentence in the middle of the screen. As soon as you read the sentence, please press any key to get the question.

If your answer to the question is YES, hit the "y" key.

If your answer is NO, hit the "n" key.

Press the "x" key to move on.

OK. Let's practice. Remember, if your answer is YES, you hit the "y" key. And if your answer is NO, hit the "n" key.

Press any key when you are ready.

Gina was eating an apple when Dick saw her.

Did Dick see Gina?

OK. Did you hit the "y" key? Good. Isn't it simple? In the experiment, you will read and answer about 40 times.

Please make sure to hit keys as soon as you can. Right after you finish reading a sentence, please hit any key. OK?

If you have any question now, please raise your hand and ask me.

If you are ready, let's start.

Press any key.

## REFERENCES

- Baker, E. L., Atwood, N. K., & Duffy, T. M. (1988). Cognitive approaches to assessing the readability. In A. Davidson & G. M. Green (Eds.), *Linguistic complexity and text comprehension: Readability issues reconsidered* (pp. 55-83). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Blau, E. K. (1982). The effect of syntax on readability for ESL students in Puerto Rico. *TESOL Quarterly*, 16, 517-528.
- Bley-Vroman, R., & Masterson, D. (1989). Reaction time as a supplement to grammaticality judgements in the investigation of second language learners' competence. *University of Hawai'i Working Papers in ESL*, 8, 207-237.
- Brown, J. D. (1996). *Testing in language programs*. Upper Saddle River, NJ: Prentice Hall.
- Brown, J. D., Yamashiro, A. D., & Ogane, E. (1999). *The emperor's new cloze: Revising cloze tests*. Unpublished manuscript.
- Brown, R. L. (1985). *A comparison of the comprehensibility of modified and unmodified ESL reading materials*. Unpublished Master's thesis, University of Hawai'i at Manoa, Honolulu.
- Chaudron, C. (1982). Vocabulary elaboration in teachers' speech to L2 learners. *Studies in Second Language Acquisition*, 4, 170-180.
- Chaudron, C. (1983). Simplification of input: Topic restatements and their effects on L2 learners' recognition and recall. *TESOL Quarterly*, 17, 437-458.

- Chaudron, C. (1985). Intake: On models and methods for discovering learners' processing of input. *Studies in Second Language Acquisition*, 7, 1-14.
- Chaudron, C. (1988). *Second language classrooms: Research on teaching and learning*. Cambridge: Cambridge University Press.
- Chung, H. (1995). *Effects of elaborative modification on second language reading comprehension and incidental vocabulary learning*. Unpublished master's thesis, University of Hawai'i at Manoa, Honolulu.
- Cohen, J., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavior Research Methods, Instruments, and Computers*, 25, 257-271.
- Dahl, H. (1979). *Word frequencies of spoken American English*. Essex, CT: Verbatim.
- Ellis, R. (1985). Teacher-pupil interaction in second language development. In S. M. Gass & C. G. Madden (Eds.), *Input in second language acquisition* (pp. 69-85). Rowley, MA: Newbury House.
- Ferguson, C. A. (1975). Toward a characterization of English foreigner talk. *Anthropological Linguistics*, 17, 1-14.
- Freedman, S. E., & Forster, K. I. (1985). The psychological status of overgenerated sentences. *Cognition*, 19, 101-131.
- Girden, E. R. (1992). *ANOVA: Repeated measures*. Newbury Park, CA: Sage.

- Green, G. M., & Olsen, M. S. (1988). Preferences for and comprehension of original and readability adapted materials. In A. Davidson & G. M. Green (Eds.), *Linguistic complexity and text comprehension: Readability issues reconsidered* (pp. 115-140). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Huckin, T., & Coady, J. (1999). Incidental vocabulary acquisition in a second language. *Studies in Second Language Acquisition*, 21, 181-193.
- Ishiguro, T. (1986). *Simplification and elaboration in foreign language teacher talk*. Unpublished Ph.D. dissertation, Stanford University, Stanford, CA.
- Issidorides, D. C., & Hulstijn, J. H. (1992). Comprehension of grammatically modified and nonmodified sentences by second language learners. *Applied Psycholinguistics*, 13, 147-171.
- Kemper, S. (1988). Inferential complexity and the readability of text. In A. Davidson & G. M. Green (Eds.), *Linguistic complexity and text comprehension: Readability issues reconsidered* (pp. 141-165). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Kim, Y. (1996). *Effects of text elaboration on intentional and incidental foreign language vocabulary learning*. Unpublished master's thesis, University of Hawai'i at Manoa, Honolulu.
- Kleifgen, J. A. (1985). Skilled variation in a kindergarten teacher's use of foreigner talk. In S. M. Gass & C. G. Madden (Eds.), *Input in second language acquisition* (pp. 59-68). Rowley, MA: Newbury House.
- Long, M. H. (1981). Input, interaction, and second-language acquisition. In H. Winitz (Ed.), *Native language and foreign language acquisition* (pp. 259-278). NY: The New York Academy of Sciences.

- Long, M. H. (1983). Native speaker/non-native speaker in the second language classroom. In M. A. Clarke & J. Handscombe (Eds.), *On TESOL '82* (pp. 207-225). Washington, D.C.: TESOL.
- Long, M. H. (1985). Input and second language acquisition theory. In S. M. Gass & C. G. Madden (Eds.), *Input in second language acquisition* (pp. 377-393). Rowley, MA: Newbury House.
- Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413-468). San Diego: Academic Press.
- Long, M. H., & Ross, S. (1993). Modifications that preserve language and content. In M. L. Tickoo (Ed.), *Simplification: Theory and application* (pp. 29-52). Singapore: SEAMEO Regional Language Centre.
- Longman. (1987). *Longman dictionary of contemporary English*. (2nd ed.). New York: Longman.
- Parker, K., & Chaudron, C. (1987). The effects of linguistic simplification and elaborative modifications on L2 comprehension. *University of Hawai'i Working Papers in ESL*, 6, 107-133.
- Pica, T., Doughty, C., & Young, R. (1986). Making input comprehensible: Do interactional modifications help? *ITL Review of Applied Linguistics*, 72, 1-25.
- Pica, T., Young, R., & Doughty, C. (1987). The impact of interaction on comprehension. *TESOL Quarterly*, 21, 737-758.

- Ross, S., Long, M. H., & Yano, Y. (1991). Simplification or elaboration?: The effects of two types of text modification on foreign language reading comprehension. *University of Hawai'i Working Papers in ESL, 10*, 1-32.
- Snow, C. E. (1972). Mother's speech to children learning language. *Child Development, 43*, 549-565.
- Tsang, W. K. (1987). *Text modification in ESL reading comprehension*. Unpublished Scholarly Paper, University of Hawai'i at Manoa, Honolulu.
- Urano, K. (1998). *Lexical simplification and elaboration: A pilot study on sentence comprehension and incidental vocabulary acquisition*. Term paper, ESL 672 (Second Language Classroom Research). Honolulu: University of Hawai'i at Manoa.
- Watanabe, Y. (1992). *Effects of increased processing in incidental learning of foreign language vocabulary*. Unpublished master's thesis, University of Hawai'i at Manoa, Honolulu.
- Wesche, M. B., & Ready, D. (1985). Foreigner talk in the university classroom. In S. M. Gass & C. G. Madden (Eds.), *Input in second language acquisition* (pp. 89-114). Rowley, MA: Newbury House.
- Yano, Y., Long, M., H., & Ross, S. (1994). The effects of simplified and elaborated texts on foreign language reading comprehension. *Language Learning, 44*, 189-219.