



Theses and Dissertations

2009-07-14

Effects of Multimedia Glossary Annotations on Incidental Vocabulary Acquisition in L2 Learners of Japanese

Brian Gleason James
Brigham Young University - Provo

Follow this and additional works at: <https://scholarsarchive.byu.edu/etd>



Part of the [Other Languages, Societies, and Cultures Commons](#)

BYU ScholarsArchive Citation

James, Brian Gleason, "Effects of Multimedia Glossary Annotations on Incidental Vocabulary Acquisition in L2 Learners of Japanese" (2009). *Theses and Dissertations*. 1812.
<https://scholarsarchive.byu.edu/etd/1812>

This Thesis is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of BYU ScholarsArchive. For more information, please contact scholarsarchive@byu.edu, ellen_amatangelo@byu.edu.

EFFECTS OF MULTIMEDIA GLOSSARY ANNOTATIONS ON
INCIDENTAL VOCABULARY ACQUISITION IN
L2 LEARNERS OF JAPANESE

by

Brian James

A thesis submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Master of Arts

Center for Language Studies

Brigham Young University

August 2009

BRIGHAM YOUNG UNIVERSITY

GRADUATE COMMITTEE APPROVAL

of a thesis submitted by

Brian James

This thesis has been read by each member of the following graduate committee and by majority vote has been found to be satisfactory.

Date

Paul Warnick, Chair

Date

Michael D. Bush

Date

Masakazu Watabe

Date

Jack Stoneman

BRIGHAM YOUNG UNIVERSITY

As chair of the candidate's graduate committee, I have read the thesis of Brian James in its final form and have found that (1) its format, citations, and bibliographical style are consistent and acceptable and fulfill university and department style requirements; (2) its illustrative materials including figures, tables, and charts are in place; and (3) the final manuscript is satisfactory to the graduate committee and is ready for submission to the university library.

Date

Paul Warnick
Chair, Graduate Committee

Accepted for the Department

Date

Michael D. Bush
Graduate Coordinator

Accepted for the College

Date

Joseph Parry
Associate Dean, College of Humanities

ABSTRACT

EFFECTS OF MULTIMEDIA GLOSSARY ANNOTATIONS ON INCIDENTAL VOCABULARY ACQUISITION IN L2 LEARNERS OF JAPANESE

Brian James

Center for Language Studies

Master of Arts

In recent years, advances in computer technology have allowed increasingly rich multimedia content to be incorporated into educational materials in many fields, including the field of language teaching. Yet as visually appealing as such products may be, we must ask whether multimedia-enriched materials actually improve learning in a measurable way. If so, individual curriculum makers can then decide whether the benefits of the multimedia materials justify the cost of purchasing and implementing them.

This study attempted to examine the effects of multimedia glossary aids on incidental vocabulary acquisition rates of L2 learners of Japanese. Subjects included 35 third- and fourth-year students of Japanese at a large private university in the United States, who read a Japanese short story using an online web application that included a multimedia-enriched glossary. A total of 27

keywords were selected from the text for inclusion in the glossary. A third were annotated with English text definitions only; another third had an English text definition plus a picture illustration; the final third had an English text definition plus a video illustration. An unannounced post-test measured vocabulary gains.

A logistic mixed models regression was performed to test for differences in acquisition rate across the annotation types. Also, due to the unique dual nature of Japanese orthography, which includes both the phonemic *kana* and so-called ideographic *kanji* characters, the regression also examined interaction between orthographic representation of the keyword and annotation type on acquisition rate.

A significant result ($p < 0.0001$) was found for annotation type as a main effect, with video-annotated words showing the highest acquisition rates. Additionally, a significant interaction ($p = 0.0139$) was observed between orthography and annotation type, indicating that multimedia glossary annotations may have affected the acquisition of phonemic *kana* representations of keywords differently than they affected ideographic *kanji* representations.

ACKNOWLEDGMENTS

I would like to thank my friends and family for encouraging me throughout a sometimes thorny path. Their support has helped me surmount challenges that would surely have overwhelmed me if I had faced them alone. Even the simplest words of kindness and advice, coming from one dear to us, can have a profound impact.

Also, I wish to express appreciation to each of my committee members: Dr. Warnick, Dr. Bush, Dr. Watabe and Dr. Stoneman, and to the Center for Language Studies Program Manager, Agnes Welch, who have all helped me navigate the new and unfamiliar path of graduate work. Without their great patience, knowledge and guidance, I could not have completed my thesis study and degree program.

Further, I would like to thank the staff of the ARCLITE Lab of the Center for Language Studies, particularly the web programmers – Dan Standage and Alan Patton. Their professionalism, hard work and expertise were instrumental in creating the application required for my study.

Finally, I am very grateful to the entire faculty and staff of the Asian and Near Eastern Languages Department, for making BYU such a wonderful place to study language and language teaching methods. The professors of the Japanese section, in particular, have been exemplars to me of excellence in teaching.

Table of Contents

Chapter 1: Introduction.....	1
Statement of Purpose.....	3
Significance of the Study.....	4
Chapter 2: Review of the Literature	6
Early Theories of Incidental Vocabulary Acquisition.....	6
Response to the Natural Approach and Input Hypothesis	8
Theoretical Support for Multimedia Learning	10
Early Experiments with Multimedia Glossing	11
Recent Experiments with Multimedia Glossing.....	14
The Unique Nature of Japanese Vocabulary Acquisition.....	18
Summary and Identifying the Current Research Gap.....	22
Research Questions.....	24
Chapter 3: Method	25
Participants	25
Treatment	26
CyberHon application.....	26
Text.....	29
Keywords.....	29
Instruments	33
Login screen.....	33
User behavior analytics.....	33
Survey.....	34
Vocabulary post-tests.....	34
Written summary.....	37
Procedure	37
Data Analysis.....	39
Chapter 4: Results	41
Research Question 1	41
Research Questions 2 and 3.....	44
Interaction between available annotation and orthography.....	46
Interaction between available annotation, viewing the target annotation and orthography.....	49
Chapter 5: Discussion and Conclusions	54
Incidental Vocabulary Acquisition.....	54
Relative Effectiveness of Multimedia Annotations.....	54
Implications	56
Limitations	59
Recommendations for Future Research.....	61
References	63
Appendix A: Informed Consent Document.....	66
Appendix B: Questionnaire	67

	viii
Appendix C: Text of Login Screen.....	68
Appendix D: Online Help Screen	69
Appendix E: Instruction Sheet	70
Appendix F: Unannounced Vocabulary Quiz Instructions	71
Appendix G: Written Summary Prompt	72

List of Tables

Table 1: Participants and Their Class Levels.....	26
Table 2: Words Selected for Glossary and their Attributes.....	32
Table 3: Incidental Acquisition Rates of Words on First Exposure by Glossary Annotation	41
Table 4: Incidental Acquisition Rates of Words on First Exposure by Orthography	43
Table 5: Model of Annotation Type and Orthography Effects on Vocabulary Acquisition.....	45
Table 6: Type 3 Tests of Fixed Effects for Annotation and Orthography	45

List of Figures

Figure 1: CyberHon screenshot featuring glossary and reading frames	28
Figure 2: Screenshot of CyberHon application during post-test activity	35
Figure 3: Graph of interaction between orthography and annotation type on proportion of vocabulary items identified correctly.....	47

Chapter 1: Introduction

In recent years, the proliferation of the personal computer and media technologies has revolutionized nearly all aspects of our lives, including our educational practices. Modern computing power allows for instructional materials featuring increasingly complex combinations of the printed word, images, audio and video. With novel presentation formats and visually appealing graphics, instructional computer software has become a mainstream part of the educational landscape. But are these multimedia tools actually more effective than traditional text-based materials? This study will investigate one specific aspect of that question.

Two pedagogical theories inform the basis of the present study. The first is Krashen's (1983) Natural Approach. Among other things, Krashen argued that learners acquire language most effectively through incidental learning rather than through intentional learning. In other words, he proposed that teachers utilize activities that give students opportunities to learn through their efforts to comprehend and use language, rather than focusing on conscious efforts to memorize and drill vocabulary, grammar patterns, and so on.

For the purposes of this study, and previous studies on which it is based, incidental learning is associated with extensive reading. In extensive reading mode, defined here as reading for overall comprehension rather than creating a word-perfect translation, readers may utilize strategies such as guessing word meanings from context to assist them. This is contrasted with intensive reading,

where the reader attempts to decode the text on a word-by-word basis, often focusing on an explicit application of grammatical rules and spending a greater amount of time on the passage than an extensive reader would. The differentiations between incidental and intentional learning, and extensive and intensive reading, are important elements in the conceptual background and methodology of this study, which will focus on multimedia aids for incidental learning.

Paivio's (1971) Dual Coding Theory is the second theory that informs the present study. Paivio asserted that the presentation method of new information creates a trace in the learner's mind. The nature of that trace depends upon the presentation mode; a verbal presentation of the information creates a verbal trace, whereas a visual presentation creates a visual trace. Further, these trace types can compound, and a combination of visual and verbal traces is theorized to produce more ready access to the learned information than a visual or verbal trace alone. This theory is an important part of the groundwork for an examination of the potential benefits of multimedia learning aids over text-only aids.

One line of empirical research that has attempted to examine the effectiveness of multimedia learning aids is the series of studies modeled on the Chun and Plass (1996) investigation of multimedia glossing. Their study was based on an investigation of instructional techniques that combined Krashen's Natural Approach philosophy with Paivio's Dual Coding Theory to investigate the effectiveness of multimedia glossary aids for incidental vocabulary learning during extensive reading tasks. The study was replicated by subsequent

researchers, who conducted studies that involved subjects who were learning various target languages. Although there was no clear consensus on which type of multimedia aid (specifically still picture versus video) was most effective, each of these researchers found that multimedia aids on the whole promoted greater incidental vocabulary acquisition than text-based aids alone. None of these researchers, however, conducted studies involving Japanese foreign language learners.

Japanese presents an opportunity for a new perspective on this established line of research. All previously researched languages have been Indo-European, with alphabetic orthographies that correspond to phonemes of the language. Japanese, however, is not Indo-European. It also features a unique dual orthography consisting of both phonemic characters and (so-called) ideographic characters. The present study explores this unique feature by comparing the effects of multimedia annotations on vocabulary acquisition across phonemic/non-phonemic orthography types.

Statement of Purpose

This study will attempt to illuminate the nature of incidental vocabulary acquisition associated with multimedia glossing for Japanese L2 learners by partially replicating the Chun and Plass (1996) landmark study. The present study, however, will address not only the same problems as previous studies, it will also address the aforementioned issue raised by the dual nature of Japanese

orthography. Specifically, this study will address the following research questions:

1. What is the rate of incidental vocabulary acquisition in Japanese L2 learners after a multimedia glossary treatment?
2. Is there a demonstrable difference between the effects of the various annotation types on vocabulary acquisition rates?
3. If vocabulary acquisition rates are different for the various annotation types, does orthography (*hiragana* versus *kanji*, i.e. phonemic versus non-phonemic characters) interact with the main effect of annotation type?

Significance of the Study

The answers to these questions have potential impact in three ways. First, it is hoped that they will help clarify the nature of previously observed multimedia effects on vocabulary acquisition in relation to Japanese, with its unique linguistic features. If the results are consistent with previous studies, it will demonstrate greater validity for those findings. Otherwise, it will suggest that previously formulated theories may need to be revised, at least in the case of Japanese.

Second, the results of this study may have import in the fields of psycholinguistics and language acquisition, as they will compare and contrast the acquisition of phonemic and non-phonemic orthographies within the same language. This may in some small way elucidate the nature of the written language itself, specifically the differences between how the mind perceives and processes phonemic and non-phonemic characters.

Third, it is hoped that these results will have practical ramifications for Japanese language teachers, who must decide what kinds of learning activities and materials are both effective and suited to their goals. Multimedia learning tools can be expensive, and their implementation requires time and effort. This study, along with others, will help empower teachers and administrators to make objective cost-benefits analyses based on numerical results rather than marketing claims.

Chapter 2: Review of the Literature

The intent of this study is to investigate the effects of multimedia glossary annotations on incidental vocabulary acquisition in L2 Japanese language learners. Before proceeding to the methodology and results of the present study, however, let us examine the background that informs the research questions at hand.

This chapter will review theoretical issues related to multimedia learning, incidental learning, and some unique characteristics of the Japanese written language. It will then describe previous empirical studies that draw on parts of this theoretical framework and are relevant to the present research questions. Finally, the chapter will conclude with a description of a gap in the current body of research literature and how this study might help fill the void.

Early Theories of Incidental Vocabulary Acquisition

In recent decades, early research of incidental vocabulary acquisition was spearheaded by those who claimed that vocabulary is best learned through exposure in a natural meaningful context, rather than through an intentional learning activity. One of these was Stephen Krashen, who, with Tracy Terrell, formulated the Natural Approach in the late 1970s. As part of the theoretical framework on which the approach is based, Krashen (1983) proposed the Input Hypothesis, which states that learners acquire language through exposure to new language forms encountered in comprehensible input in a communicative setting.

He theorized that the ideal input would be at a level just above the learner's current degree of mastery – what he termed the “*i+1*” level – and should be made comprehensible through context or extra-linguistic clues which allow students to infer the meaning of unfamiliar words and grammar for themselves. Krashen claimed that traditional classroom activities aimed at directly teaching a student vocabulary may be successful temporarily, but a word learned outside of a genuine communicative setting will never be retained in long-term memory. In short, Krashen argued that vocabulary can only be acquired in a setting where students are not attempting to directly learn the word, but are instead engaged in a communicative task which utilizes the word. Thus, he concludes, incidental vocabulary acquisition is the best way to acquire new words. Proper reading instruction, according to Krashen (1993), encourages students to read for comprehension, to infer meanings of unfamiliar words, and to strictly limit the use of dictionaries.

The Natural Approach and Input Hypothesis are supported to some extent by research in the field of first language acquisition. Particularly of interest is the 1985 study by Nagy, Herman and Anderson. A group of middle school students was divided into two groups, with each group reading either a narrative or an expository text. The students were instructed to read for comprehension, and were given an unannounced multiple-choice vocabulary post-test after the reading.

According to the findings, after only one exposure to a new word students had a 10% - 15% chance of acquiring any given unfamiliar word from the reading well enough to identify it on the post-test. This figure may at first seem too low

to be of consequence. Nagy, Herman and Anderson argue, however, that since students are exposed to hundreds of thousands of words per year, even at a modest acquisition rate of 10%, they could potentially learn thousands of new words. In contrast, direct instruction in vocabulary can treat only a limited number of items per session, and may not help students become independent vocabulary learners outside the classroom. The researchers concluded that while direct vocabulary instruction may have its place in some contexts, encouraging incidental vocabulary acquisition through reading is more effective at building vocabulary.

Response to the Natural Approach and Input Hypothesis

Later studies by second language researchers, however, have shown that a reading strategy involving vocabulary lookup may also be beneficial, at least for L2 learners. Knight's (1994) study investigated the effect of using a computerized dictionary on vocabulary acquisition by 105 L2 Spanish learners. Subjects were stratified by verbal ability, as measured by their American College Test verbal scores, and were randomly placed into either a no-dictionary group or a dictionary access group, and were asked to read several short articles from a Spanish language magazine.

Those who had access to the dictionary while reading performed significantly better on both immediate and delayed vocabulary post-tests than learners of the same verbal ability in the no-dictionary group. In other words, both high and low verbal ability readers benefited from the dictionary treatment. Dictionary use appeared to be of particular utility to the low verbal ability subjects, as it tended to reduce the difference between their scores and the scores

of high verbal ability subjects. Knight concluded that, although it is possible for students to acquire new vocabulary by inferring meaning, dictionary use enhances the acquisition process overall, and helps to level the playing field for low verbal ability learners. She called for language teachers to reconsider the then-common practice of encouraging students to guess at word meanings rather than look them up.

Seeing some evidence of the benefits of word lookup for vocabulary acquisition, Hulstijn, Hollander and Greidanus (1996) asked whether looking up words in a marginal gloss would be more effective than using a dictionary. They compared incidental vocabulary acquisition among seventy-eight students of French in three experimental groups: users of printed marginal glosses, users of dictionaries, and a control group with no reference materials. Each group was asked to read the same French short story and later took a vocabulary post-test. Test words were carefully selected to be unfamiliar to the subjects, and post-test results confirmed a negligible degree of prior knowledge of the test items for each group. Under these conditions, the marginal gloss group clearly outperformed both the dictionary group and the control group. This may be due to the fact that dictionary users looked up fewer words than did the marginal gloss users, although when a dictionary user did look up a word, there was a higher probability that he or she would remember it on the post-test compared with the marginal gloss users.

Hulstijn, Hollander and Greidanus' experimental evidence showed that looking up unfamiliar words, either in a dictionary or in a marginal gloss, facilitates the vocabulary acquisition process. Yet they also observed that

students will rarely go to the trouble of looking up those words in dictionaries. Thus, they recommended making the lookup process more convenient for readers, as they did in this study with marginal glosses, in order to encourage students to look up more words and thereby increase vocabulary acquisition.

Theoretical Support for Multimedia Learning

As the potential to improve incidental vocabulary acquisition through marginal glosses and dictionaries became clearer, computer technology was simultaneously developing in ways that would offer exciting new possibilities to enhance the lookup process through multimedia content. Glosses that describe or define a word in more than one modality, such as text and image, or sound and video, became a viable possibility for curriculum developers. Beginning in the mid-1990s, more and more research was undertaken to examine the potential benefits of multimedia glossing.

Much of the theory supporting the superiority of multimedia glossing over traditional text-only glossing comes from Paivio's (1971) work on Dual Coding Theory. Paivio framed his discussion within the historical debate between theories which posit images as mediating access to ideas in the mind versus those which contend that ideas are accessed through association with a word. He argued that both models of encoding are partially correct: the mind can store information by means of either or both types of association, but with different effects depending on the type of information being stored. Based on experimental evidence, Paivio argued that image encoding effectively facilitates recall of concrete objects, but is less effective for recalling abstract notions.

Conversely, verbal encoding is more effective for the recall of abstract notions, but less effective for recall of concrete objects.

While the whole of Paivio's work is too broad and intricate to cover within the scope of this study, two further points are especially relevant to the topic of multimedia glossing: (1) While imagery may prove more effective in some contexts, verbal encoding is also necessary for any learning activity that involves a verbal task, such as reading. (2) Although both encoding systems are distinct, they also interact, and a greater number of encoding traces (e.g. picture and text as opposed to text alone) tends to enhance the ability of the mind to recall an idea. The predicted benefits of multimodal glossing are clear. Yet although the Dual Coding Theory had existed for some time, its application to glossing had to wait until computers became powerful enough to handle high-quality multimedia content.

Early Experiments with Multimedia Glossing

In 1993, Lyman-Hager, Davis, Burnett and Chennault (as cited in Al-Seghayer, 2001) conducted an early study on the effects of multimedia glossing on L2 French vocabulary acquisition. Two groups of subjects participated in the research. Each group was asked to read the same text, but one of these groups read the text via a computer program with a multimedia gloss, while another group read from a printed text with print glosses. After one week, the computerized multimedia group demonstrated a greater retention of key vocabulary terms during an in-class discussion.

Chun and Plass (1996) conducted a series of three studies of multimedia glossing involving university students of German. More than ten years later,

research in this field seems largely based on their 1996 methodology. They created a computer program that combined the text of a German short story with a multimedia sidebar gloss to allow simultaneous glossing and reading. They selected certain key words to include in the gloss, then annotated them in one of three ways: (a) text only, (b) text and still picture, or (c) text and video. Note that words thus annotated shall hereafter be called “text-only words,” “text-plus-picture words” and “text-plus-video words,” respectively. Subjects were allowed to freely look up any of the annotation types available, or not use the gloss at all. They were told to read for comprehension, but a surprise vocabulary quiz was administered after the reading in order to measure incidental vocabulary acquisition. By comparing the vocabulary quiz scores (corrected for prior knowledge) with the number and type of annotations used by the subjects, Chun and Plass attempted to determine which mode of glossing was most effective for incidental vocabulary acquisition.

Chun and Plass do not report any significant differences between annotation types in the first study of the series, although this was treated as a pilot test. In the second study of the series, which included the largest number of subjects ($n=103$) and which the researchers themselves seemed to favor as the strongest of the three, they found that text-plus-picture words showed the greatest recall rates, followed by text-plus-video words and finally text-only words. Although it was expected that multimedia-annotated words would be more frequently remembered on the post-test, Chun and Plass noted that recall rates for text-plus-picture words were significantly higher than those of text-plus-video words, and recommended that the nature of this difference should be

investigated. A second, delayed vocabulary post-test showed an even greater increase in recall for the text-plus-picture words. Chun and Plass speculated that this may be due to “the so-called hypermnesia effect, which predicts better recall of pictures over time, compared to words” (p. 193). Their research seemed to show not only the greater effectiveness of a dual-coded glossary annotation over text-only annotations, but also the possibility of a hypermnesia phenomenon which increases long-term retention of newly acquired vocabulary.

Yoshii and Flaitz’s (2002) experimental results also support the Dual Coding Theory. Their study investigated the effects of gloss annotation types on the vocabulary acquisition of ESL students as they read a short story from a multimedia program. One group of students was given text-only annotations, while the second group was given picture-only annotations and the third group received combined text-plus-picture annotations. As predicted by the Dual Coding Theory, the combination group outperformed the other two groups on vocabulary tests of picture recognition, word recognition and providing definitions in both immediate and delayed vocabulary tests. However, delayed test results showed an equal decline in the scores of members of every group. So while Yoshii and Flaitz’s study corroborated the Dual Coding Theory and was consistent with the results of previous researchers, it did not provide any evidence to support the hypermnesia effect.

Recent Experiments with Multimedia Glossing

Subsequent researchers have also found that multimedia glossing promotes incidental vocabulary acquisition. In addition to the multimedia-versus-text question, however, these studies often included research questions which explicitly addressed the relative effectiveness of different types of multimedia annotations. Iheanacho (1997) found that both picture and video annotations in multimedia glosses were effective treatments to promote incidental vocabulary acquisition for students of English. However, in contrast to Chun and Plass, he found that words with video annotations were recalled slightly better than words with still picture annotations in a delayed post-test.

Al-Seghayer (2001) also studied the effects of multimedia glosses on vocabulary learning among ESL students. In a study similar in methodology to Chun and Plass, he created an electronic text with a gloss of words annotated with either text only, text-plus-picture, or text-plus-video. The program was administered to thirty ESL students, who were instructed to read for comprehension.

Also in contrast with Chun and Plass, Al-Seghayer found that text-plus-video words were the most frequently remembered, with students recalling an average of 6.1 out of 7 words thus annotated. This was followed by text-plus-picture words, at 4.7 out of 7 words. Text-only words were remembered slightly less often, at an average of 4.03. Despite the variance with Chun and Plass (1996) regarding the relative effectiveness of picture versus video annotation, Al-Seghayer's study coincided with Chun and Plass in his finding that both modes of multimedia annotation surpassed the text-only annotations.

In addition to the unannounced vocabulary quiz after the treatment, Al-Seghayer administered questionnaires and interviews to collect qualitative feedback from the subjects regarding the perceived degree of helpfulness of the various annotation types. The results of the questionnaires and interviews coincided with the quantitative results, indicating that students generally found video annotations most helpful in remembering new words from the text, followed by pictures, with text-only annotations least preferred.

Al-Seghayer proposes two possible explanations for his results. First, moving images may be inherently more suited to forging a connection between the linguistic form of a new vocabulary item and the meaning associated with it. Alternatively, video annotations may have simply been more interesting to watch, thus engendering a more positive affective response from the students and encouraging them to pay more attention to the word and its annotation.

It would appear that many researchers have been convinced of the effectiveness of multimedia glossing in relation to incidental vocabulary acquisition, since the research questions on the subject have more recently moved away from a basic defense of multimedia glossing to more sophisticated issues such as individual learner variables. For example, Abraham's (2001) study of Spanish language learners was similar in structure to preceding ones, but with a new hypothesis.

Going beyond previous research questions, his electronic reader offered text, video and picture annotations for all words in the glossary, and attempted to discover the effects of cognitive load by comparing free-glossing in a multimedia environment, where students are allowed to choose which

annotations to view, versus forced glossing, where students had to view all annotations available for every word in the glossary. A third control group was included, which read the same computerized text but did not have any gloss available. He also looked for a relationship between (a) subjects' spatial reasoning as measured by the Paper Folding and Card Rotations Tests (Ekstrom, French, Harman and Derman, 1976 cited in Abraham, 2001), and verbal ability as measured by the Advanced Vocabulary Test (Ekstrom, French, Harman and Derman, 1976 cited in Abraham, 2001) and (b) the apparent effectiveness of multimedia annotations. Finally, after the treatment, he administered surveys to the participants to discover their preferences among annotation types.

The results of Abraham's experiment are interesting. First, questionnaire responses indicated that 64.7% of free-lookup subjects found picture annotations useful for vocabulary acquisition, and the same number found video annotations useful. Of the forced-lookup subjects, 81.9% found picture annotations useful, and 78.8% found video annotations useful. Second, while vocabulary pretests showed no significant differences between the control, free-lookup and forced-lookup groups, the two lookup groups significantly outperformed the control group on the post-test. The control group scored an average of 6.91 out of a possible 20 points on the test, while the free-lookup and forced-lookup groups scored an average of 9.38 and 9.71, respectively.

The difference between lookup and non-lookup groups was expected, but it is interesting to note that the forced-lookup group marginally outperformed the free-lookup group. Since the forced-lookup group had to view each piece of annotation data, the potential existed for less spatially-acute members of the

forced-lookup group to suffer from cognitive overload as a result of undesired extraneous visual input. This did not appear to be the case, however, and further analysis showed that neither spatial nor verbal abilities appeared to affect vocabulary gains. Abraham is careful to acknowledge methodological weaknesses which may have prevented such an effect from appearing, but Chun and Payne (2004) would find a similarly surprising result in a later study.

Chun and Payne (2004), noting both the importance of individual learner attributes and the possibility that cognitive overload from multimedia annotations may detract from some learners' incidental vocabulary acquisition, conducted an experiment similar in concept to Abraham's. This experiment, however, would look at working memory and gloss use. They measured subjects' verbal working memory through a nonsense word repetition test, and through a test to determine reading span, then classified subjects' working memory as "high" or "low" according to each test. Every subject read a German short story using the same multimedia program Chun and Plass created for their 1996 study. After reading, the subjects were asked to complete a vocabulary quiz.

Chun and Payne found no correlation between working memory and vocabulary quiz scores. It appeared that students with limited working memory tended to acquire just as much vocabulary as their counterparts with greater working memory. While this may seem counterintuitive, Chun and Payne noted that low working memory students also tended to look up words more frequently. They speculated that low working memory subjects used the gloss to the point that they were able to compensate for their limited working memory

and were able to perform at the same level as high working memory subjects on the vocabulary quiz. Although they acknowledge the small participant pool for this study (thirteen subjects), Chun and Payne's work reflects the trend toward more sophisticated studies into the application and potential benefits of using multimedia glosses for vocabulary acquisition.

While research has been conducted in this area involving learners of Spanish, French, German and English, there seems to be no research on the effects of multimedia annotations on incidental vocabulary acquisition of L2 learners studying languages with a non-phonographic orthography, such as Japanese or Chinese. This is significant because Japanese, along with Chinese, which is the basis for much of Japanese orthography, presents a new twist on some of the fundamental questions regarding the effectiveness of multimedia annotations in incidental vocabulary acquisition through reading.

The Unique Nature of Japanese Vocabulary Acquisition

Japanese is unique to this field of study, first and most obviously because it exists outside the Indo-European family, which is the predominant language group studied by multimedia vocabulary acquisition researchers. Second, and perhaps more importantly, Japanese has a unique orthography the likes of which has yet to be explored by the current research.

Unlike the languages involved in the research cited here, Japanese words can be written in any of three independent character sets: *hiragana*, *katakana* or *kanji*.¹ The first two, collectively known as the *kana* characters, are phonemic orthographies comprised of characters corresponding to each

¹ For very specific purposes, Roman alphabet characters, or *romaji*, are also used by native

phoneme of the spoken language, where each character represents a vowel, a consonant + vowel combination, or the nasal phoneme /n/. The *hiragana* and *katakana* are identical in phonemic content and essentially differ only in form. Japanese writers choose which character set to use based on pragmatic and stylistic conventions. While some whole words in modern prose are written in *kana*, the use of this script tends to be limited to the expression of grammatical words and inflections (*hiragana*), and foreign loan words (*katakana*).

The bulk of content words in Japanese is written in *kanji*. *Kanji* characters were borrowed from Chinese over the course of several eras in Japan's ancient history. Unlike the *hiragana* and *katakana*, the *kanji* do not always have readily apparent phonemic values, making their pronunciation somewhat more opaque than that of the *kana*. Each *kanji* represents a different word or concept, and in most cases, pronunciation must be acquired on a character-by-character basis. It must be noted that this phonemic opacity is not perfect, as there are groups of *kanji* that share structural similarities, providing clues to their pronunciation. Suffice it to say, however, that the *kanji*, on the whole, could be considered phonemically opaque relative to the *kana*, or to the alphabetic orthographies of the languages targeted in the studies discussed above. Thus, the traditional view of *kanji* places them in the category of ideographs – characters that express meaning as images, without recourse to the spoken language and its phonemic elements.

Japanese to express Japanese words. However, the use of *romaji* is narrow, and typically limited to graphic design applications such as environmental signage, corporate branding and advertisements. Since native Japanese typically do not extensively use *romaji* to express Japanese words in their literature, the present research does not consider *romaji* a significant element of Japanese orthography for vocabulary acquisition through reading.

However, a number of researchers (DeFrancis 1984, Matsunaga 1995, Everson and Ke 1997, Everson 1998, Mori 1998, Erbaugh 2002, Unger 2004) have shown that *kanji* characters are not simply ideographs. The conclusions of that debate and the ultimate choice of classification for *kanji* characters are not germane to the present research, but the debate touches on issues that are of relevance here, and these particular points will be discussed below.

Researchers such as DeFrancis (1984) and Unger (2004) have made the case against an ideographic classification of Chinese characters based on theoretical grounds, arguing that all orthographies have reference to the spoken language, since spoken language is linguistically primary to writing. Therefore, they contend, all orthographies are essentially phonemic in nature, even if their phonemic values tend towards opacity, like the *kanji*. Empirical psycholinguistic research by Matsunaga (1995) suggests that native Japanese readers recognize *kanji* characters at a phonemic level rather than as ideographs. Further studies by other researchers have illuminated the connection between *kanji* characters and phonemic processing, not only among native speakers, but among adult language learners as well.

Language acquisition researchers studying Chinese foreign language learners have found similar results. Everson and Ke's (1997) qualitative study of seven intermediate and advanced university students used verbal report instruments, also known as *think aloud protocols*, to investigate subjects' reading strategies. Subjects were given newspaper passages to read, while the researchers noted both their verbal reports and other observable reading behaviors. Among several findings, it was discovered that subjects exhibited

“countless instances” of sound remediation, such as subvocalization and sounding out characters while reading. This suggests a connection between reading *kanji* and phonemic processing, which would be further examined in later research.

Everson (1998) subsequently conducted a study aimed specifically at examining this connection. Subjects consisted of twenty first-year students of Chinese as a foreign language. Each subject was shown 46 *kanji* characters that had already been introduced in the university’s first-year Chinese curriculum; subjects were then tested on both pronunciation and translation for each character. Analysis indicated that pronunciation ability was highly correlated with translation ability. Perhaps more tellingly, Everson also found that for any given word that was pronounced correctly, there was a 90.7% chance of the subject correctly identifying it. If the word was incorrectly pronounced, however, there was only a 12.0% chance of correctly identifying it.

Thus the ability to pronounce a *kanji* character and the ability to grasp its meaning appear to be linked not only in a general sense, but also on a word-for-word basis: if an individual knows how to pronounce a character, he or she is highly likely to be able to recall its meaning as well. Conversely, if he or she is unable to produce a character’s sounds, the individual is unlikely to know its meaning.

Aoyama (2005) found consistent results in his study with a Japanese diglot reader. He administered a computer-based reading treatment to 81 first-year university students of Japanese, where the subjects were asked to read a prose selection primarily in English, with keywords in the target language embedded

into the body of the text. A gloss was available in the reading application, which included differing pronunciation helps for different words. Aoyama found that out of all the *kanji* that subjects learned through his treatment, the majority – 62 percent – were learned as a combination of meaning and pronunciation, while only 38 percent were learned with meaning only. Further, a delayed post-test found a significantly higher retention rate of *kanji* learned as pronunciation and meaning, as opposed to *kanji* learned by meaning alone.

Similar findings have been made by other researchers, which highlight the connection between mastery of pronunciation and meaning of a *kanji* character. This evidence prompted Koda (2005), in her review of second language reading research, to conclude that phonological decoding is the most important ability for reading in any foreign language, regardless of orthography. It seems well-established that phonemic acquisition of a *kanji* character is extremely beneficial or perhaps even critical to learning a *kanji* character and its meaning. However, no research has been conducted on visual multimedia glossary annotations and their relative effects on the acquisition of phonemically opaque vocabulary representations such as *kanji*, as opposed to phonological representations such as the *kana*. Japanese presents a unique opportunity to do this, since most words can be represented in either of the two orthographies.

Summary and Identifying the Current Research Gap

Theory and experimental evidence regarding incidental vocabulary acquisition and multimedia glossing have come a long way since the dawn of the Natural Approach and Dual Coding Theory. Krashen's Input Hypothesis discouraged glossing during an extensive reading activity, yet later research

indicated that frequent glossing during extensive reading may aid vocabulary acquisition. Not long afterward, computer technology had progressed to the point that multimedia annotations became a possibility.

As early as 1993, researchers began to discover the potential of computer-based multimedia instruction. Early experiments in the field were relatively simple comparisons of traditional text-based glossing with multimedia glosses. But once the benefits of multimedia glossaries had been established, more rigorous research was devised to explore the interaction of various individual learner factors with multimedia effects. A number of interesting findings and hypotheses have been put forth as researchers attempt to improve our use of multimedia glosses as a tool for incidental vocabulary acquisition.

A second strand of research investigated the nature of the *kanji* characters, which are utilized to express much of the vocabulary in both Chinese and Japanese. The work of those who contest the assertion that *kanji* are ideographs, as well as the findings of language acquisition researchers, have illuminated an interesting connection between learning the pronunciation of a *kanji* and learning its meaning. Mastering pronunciation seems to play an important role in the semantic acquisition of vocabulary, including vocabulary expressed in *kanji*.

These two strands of research will be tied together in the present study to answer novel questions. Largely due to its orthography, reading Japanese is quite unlike reading the languages discussed in previous multimedia glossing research. Japanese has both a native phonemic character set – the *kana* – as well as a phonemically-opaque character set – the *kanji*. This dual orthography will allow us to examine not only the effect of multimedia glossing on incidental

vocabulary acquisition, but also the relative effects of multimedia glossary annotations on the acquisition of both phonemic and so-called ideographic representations of vocabulary.

Research Questions

The research gap mentioned above suggests the following research questions:

1. What is the rate of incidental vocabulary acquisition in Japanese L2 learners after a multimedia glossary treatment?
2. Is there a demonstrable difference between the effects of the various annotation types on vocabulary acquisition rates?
3. If vocabulary acquisition rates are different for the various annotation types, does orthography (*hiragana* versus *kanji*, i.e. phonemic versus non-phonemic characters) interact with the main effect of annotation type?

Chapter 3: Method

Adapting the methodology of Chun and Plass (1996) to the research questions discussed in Chapters 1 and 2, the present study solicited volunteer subjects from Japanese classes at a large university and administered an experimental reading treatment. The subjects were asked to read a short story in Japanese on a computer application that displayed both the text of the story and an on-screen multimedia glossary. A third of the words in the glossary were annotated with text translations but no multimedia content. Another third of the words were annotated with text translations and an illustrative picture. The final third were annotated with text translations and an illustrative video. Two unannounced post-tests measured subjects' acquisition of vocabulary from the treatment. The first test measured recognition of the *kanji* representation of each vocabulary item included in the glossary. The second test also tested recognition of each vocabulary item, but displayed the keywords in *kana* form. The results were subjected to statistical analysis in an attempt to answer the research questions. The following subsections describe various aspects of the present methodology in detail.

Participants

Subjects in this study consisted of 35 students of Japanese at Brigham Young University who were currently enrolled in a third- or fourth-year Japanese reading class at the time of the study. Volunteer subjects were solicited from the following four courses during the Fall 2008 and Winter 2009 semesters: Japanese 321 – Reading Modern Documentary Styles; Japanese 322 – Reading

Modern Short Fiction; Japanese 443 – Advanced Documentary Style; Japanese 444 – Reading Modern Japan. Due to the demographic makeup of the university, it was conjectured beforehand that a large number of subjects may have participated in a one-and-a-half to two-year volunteer program in Japan. The initial methodology called for soliciting only subjects who had participated in this program in order to establish a homogeneous subject pool. It soon became clear, however, that this stipulation would reduce the subjects to an unacceptably low number. Thus, volunteers were solicited regardless of prior participation in that program.

Table 1

Participants and Their Class Levels

Japan 321	Japan 322	Japan 443	Japan 444	Total
15	10	2	8	35

* *Class levels are counted as the highest concurrently enrolled Japanese course*

Treatment

CyberHon application. The researcher, in collaboration with programmers in the ARCLITE Lab of the BYU Center for Language Studies, developed an online web application to administer the experimental treatment for this study. The application was dubbed *CyberHon*, in honor of the original Chun and Plass (1996) *CyberBuch* computer program.² The *CyberHon* application was hosted on university web servers and was available for use by subjects during the data gathering phase of the study, as described below in the Procedures section.

² *Buch* is German for “book,” while *Hon* is the Japanese equivalent.

CyberHon is an online reading environment that displays a Japanese short story in one frame of the screen while displaying a multimedia gloss in another frame on the same screen. Selected keywords are hyperlinked in the body of the text, and when clicked, they display content in the glossary frame for that particular word. Specifically, when a keyword is clicked, the glossary first displays both the *kanji* and *kana* representations of the keyword. This is automatic and happens whenever a keyword is clicked. From there, the subject must click buttons at the top of the screen to access the desired annotations, such as text, picture or video, that are available for that particular word.

For any given keyword, the availability of annotation content is determined by the annotation category into which the keyword has been placed. As described above, each word in the glossary falls under one of three categories: text-only, text-plus-picture or text-plus-video. Note that for each keyword, regardless of category, a written English equivalent, as well as an audio recording of a native speaker pronouncing the word is available. Text-only words have only that verbal annotation content. The other categories have multimedia annotations, *in addition to* the verbal content. Thus, text-plus-picture words have the verbal content described above, as well as a still picture illustrating the keyword. Similarly, text-plus-video words have verbal content as well as a silent video that depicts the keyword. The process of selecting and placing keywords into annotation categories is described in the subsections below.



Figure 1. CyberHon screenshot featuring glossary and reading frames.

Using *CyberHon*, subjects can read the short story page by page, and are allowed to navigate forward and backward through the story at will. A simple help screen is available at any time during the reading session, and instruction screens are automatically displayed as the subject progresses from the reading to the post-reading activities, and between each post-reading activity. A more detailed description of the post-reading activities is included in the *Instruments* subsection below.

Text. The text for the reading treatment is the short fiction piece *Naruhodo* by Hoshi Shinichi. This story was chosen based on two main criteria: (1) The researcher projected that the story's length would allow it to be read in less than an hour by students at the proficiency level relevant to this study; and further, (2) the researcher also determined that the story's vocabulary content would be challenging enough to provide unfamiliar words that could be learned through the treatment, while remaining simple enough to allow subjects to complete the reading in a timely manner. The text was divided into thirteen pages for online display in *CyberHon*.

Keywords. A total of 27 keywords were selected from the text for inclusion in the glossary. These were split into three groups of nine; each group of words comprised one of the three annotation categories discussed above. To reiterate, words in the first category, designated as "text-only," were annotated with an English text definition of the Japanese word from the story. Words in the second category, or "text-plus-picture" group, were each annotated with an English text definition and a still picture representation of the word. Words in the third category, "text-plus-video," were annotated with an English text definition and a short, silent video representing the word. Every item in the glossary also featured an audio recording of a native speaker saying the word as a pronunciation aid.

Special care was taken to ensure homogeneity between groups of words during the keyword selection process. The first condition was that any keyword included in this study must be written in *kanji*, as opposed to *kana*, in the text of the story. *Kana* exposure to the keyword was designed to come from the

glossary, when the keyword was clicked. This was done in an attempt to ensure a uniform exposure to the *kanji* and *kana* representations of each word in terms of glossary versus text encounters.

The second condition was that there should be a reasonable chance that an intermediate or advanced level student might not know the keyword already, since vocabulary acquisition could not be measured if every keyword were already known before the treatment. Thus, extremely simple words were excluded. Deciding which words fit this criterion was a subjective judgment on the part of the researcher, but was not wholly uninformed, as the researcher recently graduated from the same Japanese program at the same university as the subjects.

The remaining glossary candidates were evaluated in terms of six criteria: type of word (action, object or description), subtype (human or non-human for action and description, abstract or concrete for object), difficulty (easy or hard), frequency in the present text, importance in the clause in which it appears (very important or less important), and page number from the original text on which the word occurs. Again, many of these characteristics were objective and readily determined, such as type, subtype, frequency and page number. Yet the other two, difficulty and importance, were of necessity subjective judgments made on the part of the researcher.

The resultant number of 27 keywords was the result of the two general principles governing word selection: The researcher attempted to include as many words as possible to generate the greatest amount of data possible, while assuring homogeneity of the words comprising the annotation categories.

Ultimately, a total of 27 words were found that could be blocked out into roughly equivalent groups. Going beyond this number with the present text, given the criteria for homogeneity discussed above, would have created a degree of heterogeneity that, in the judgment of the researcher, would have confounded any comparison of vocabulary gains across different annotation categories. Thus, 27 was the maximum number of keywords that were available in the text, that could be arranged into sufficiently homogeneous groups. The final selection of keywords and their groupings are given in Table 2.

Each group of words was arbitrarily assigned to an annotation type. Group 1 was annotated with text only; Group 2 was assigned text plus picture; Group 3 included text plus video. As per the design of this study, no words were annotated with both picture and video.

Note that some of the keywords in the text-plus-picture or text-plus-video categories may not appear to be suitable for annotation with picture or with video. The attribute of *suitability for multimedia portrayal* was intentionally left out of consideration when choosing and assigning keywords to an annotation category. This was done for three reasons. First, measuring suitability for multimedia portrayal is highly subjective, even more so than word difficulty or importance in a clause. Measurements of the latter two, unlike the former, can be defended or challenged at least in part through recourse to rational argument. Second, words that share suitability for a particular multimedia portrayal may

Table 2

Words Selected for Glossary and their Attributes

Group	Item	English	Type	Subtype	Difficult	Freq	Import	Page*
1	連行される	To be led (away)	ACT	HUM	Easier	1	Lesser	15
1	主張する	To claim	ACT	HUM	Harder	1	Lesser	17
1	一刻も早く	Immediately	DESC	HUM	Easier	2	Lesser	17
1	試験	Test	OBJ	CON	Easier	1	Lesser	14
1	歓声	Cheer	OBJ	CON	Easier	1	Lesser	18
1	報告	Report	OBJ	CON	Harder	1	Lesser	17
1	留置場	Jail	OBJ	CON	Easier	2	Lesser	15
1	警官	Police officer	OBJ	CON	Harder	3	Greater	14
1	宇宙人	Space alien	OBJ	CON	Harder	8	Greater	16
2	到着する	To arrive	ACT	HUM	Easier	1	Lesser	14
2	控訴する	To appeal (legal)	ACT	HUM	Harder	1	Lesser	18
2	無罪	Not guilty	DESC	HUM	Easier	3	Greater	16
2	参考書	Notebook	OBJ	CON	Easier	1	Lesser	14
2	うそ発見機	Polygraph	OBJ	CON	Easier	1	Lesser	17
2	円盤	Saucer	OBJ	CON	Harder	1	Lesser	16
2	住居	Home	OBJ	CON	Easier	2	Lesser	16
2	検事	Prosecutor	OBJ	CON	Harder	4	Greater	15
2	被告	Defendant	OBJ	CON	Harder	7	Greater	15
3	分解する	To disassemble	ACT	HUM	Easier	1	Lesser	16
3	混乱させる	To confuse	ACT	HUM	Harder	1	Lesser	16
3	有罪	Guilty	DESC	HUM	Easier	3	Greater	17
3	置時計	Table clock	OBJ	CON	Easier	1	Lesser	14
3	消火器	Fire extinguisher	OBJ	CON	Easier	1	Lesser	15
3	依頼人	Client	OBJ	CON	Harder	1	Lesser	18
3	交渉	Conference	OBJ	CON	Easier	2	Lesser	18
3	惑星	Planet	OBJ	CON	Harder	3	Greater	16
3	犯人	Criminal	OBJ	CON	Harder	9	Greater	15

* Measured by where the keyword appeared in the original printed version, not in *CyberHon*.

also share other features, such as ease of visualizing, that could create uncontrolled bias if they were to be intentionally grouped together into one annotation category. Third, the present study addresses the question of

multimedia aids and their effectiveness in promoting incidental acquisition of vocabulary in a general sense, not the acquisition of vocabulary that is particularly suitable to multimedia presentation. It is a given, then, that some of the words may not appear particularly suited to portrayal in multimedia. Indeed, it would be expected that some words should not be particularly so suited. For these reasons, the researcher found it problematic to use suitability for multimedia portrayal as a criterion in the keyword selection process, and elected not to do so.

Instruments

Login screen. *CyberHon* acts not only as a treatment, but also as a data gathering instrument. Data gathering begins with the initial login screen, where subjects are asked to enter their full names, student ID numbers and E-mail addresses. This information is used to distinguish unique users.

User behavior analytics. Another of the data gathering functions incorporated into *CyberHon* is a routine to record user behavior. The following user actions are logged for each subject as he or she reads the short story: (1) loading a new page, i.e. when the user “turns” the online page; (2) when a keyword from the text is clicked; (3) when a glossary annotation, such as picture or text, is accessed. These data, particularly the glossary access data, create a more complete picture of the user activity and reading experience, and were intended to assist in statistical analysis of the results. Comparison of vocabulary acquisition rates with the mere *availability* of multimedia content is weaker in terms of validity than comparing vocabulary acquisition rates with the actual

viewing of glossary content. *CyberHon's* behavior recording functionality allows the necessary data to be gathered for the latter, stronger analysis.

Survey. A paper-based questionnaire was devised with the intent of gathering biographical data relative to the subjects' background studying the Japanese language. The questionnaire asks subjects to indicate prior Japanese experience and currently enrolled Japanese classes. The survey can be found in Appendix B.

Vocabulary post-tests. Once again, *CyberHon* is utilized as a data gathering instrument on several counts, including vocabulary acquisition. Two unannounced vocabulary post-tests are incorporated into the application, to be administered after the reading session. The first one tests recognition of each of the 27 *kanji* keywords from the text. The second one tests recognition of the phonemic *kana* representation of every keyword. Thus, every keyword is tested twice, once in its *kanji* form, and once in its *kana* form. Both of the tests are in multiple-choice format, with the English equivalent of a keyword given as a prompt, followed by six possible Japanese equivalents, in *kanji* during the first test, and *kana* during the second, underneath. The order of the answer choices and the order of the prompts were each randomized once and then presented in that order to each student.

The multiple-choice, recognition format of the post-tests in the present study is similar to the format of the post-test in the third and final study in Chun and Plass's 1996 report. While their first two studies employed tests of



Figure 2. Screenshot of CyberHon application during post-test activity.

vocabulary production, Chun and Plass ultimately decided to utilize a multiple-choice, recognition test in the final study. They explained that this format would better fit the nature of the treatment, which involved incidental learning and vocabulary recognition rather than production. Note that Chun and Plass's recognition test used the target annotation of each keyword as a test prompt: an English text definition for text-only words, the picture from the glossary for text-plus-picture words, and the video from the glossary for text-plus-video words. The present study, however, uses only English text definitions of the keywords as prompts for every item, regardless of annotation.

This divergence from the Chun and Plass methodology brings the present post-test into further agreement with the purpose of the experimental treatment. The ultimate objective of the *CyberHon* treatment, if it were deployed in classrooms, would be to increase subjects' verbal recognition of target language words, not their ability to associate pictures or videos with target language vocabulary. Further, using picture or video prompts that are unique to the present treatment might not accurately reflect how well subjects could apply their learning from *CyberHon* to other reading contexts. Again, the purpose of the treatment in a "real" classroom setting would not be to increase subjects' ability to read this particular text with these particular multimedia aids, but rather to increase overall reading vocabulary.

The six possible answers for any given test item consisted of the correct answer and five distracters that were chosen randomly from the pool of other keywords in the text. The researcher decided on this course for two reasons: First, it ensures a more equal difficulty between test items, rather than intuitively choosing distracters for each item with similar pronunciation or, in the case of *kanji*, similar radicals or other structural components. Second, although the distracters do not necessarily bear close resemblance to the correct choice, both the correct choice and the distracters would have been recently read, and probably glossed, by the subject during the reading treatment. The researcher hoped that this would provide enough of a similarity to make the other choices effective distracters for a subject who was unsure of the correct choice.

In addition to the prompt and multiple-choice answers, a drop-down box was included with options for the student to indicate which glossary annotation

most readily came to mind when trying to answer the test item. One of the options in this box, labeled “Already Knew,” was provided to gather information that would help control for subjects’ knowledge prior to the treatment.

Written summary. To keep subjects focused on the task in an extensive reading mode, the researcher informed subjects before they started the reading session that they would be asked to write a brief English summary of the text after finishing the story. Once more, *CyberHon* was utilized to record subject responses to this activity, although the data was not intended for analysis. For the purposes of the present study, the written summary was essentially presented to preserve the face validity of the treatment.

Procedure

Subjects were recruited in two phases: a smaller group in Fall Semester of 2008 from Japanese 321 and 322 courses at Brigham Young University, and a larger group in Winter Semester of 2009 from Japanese 321, 322, 443 and 444.

In each phase of the research, subjects were asked to come to an on-campus computer laboratory which had been reserved for the study. The laboratory was equipped with multiple Apple Intel-powered iMac computers running Macintosh OS X. After reading an informed consent form (Appendix A), subjects were asked to complete a printed survey (Appendix B). Next they were asked to open the Safari web browser³ and navigate to the URL of the *CyberHon* login page. After entering their personal information and reading the login screen instructions (Appendix C), students were presented with the main

³ *CyberHon* appears to be compatible with Firefox, but was coded specifically for Safari.

CyberHon reading page, including the story frame and the glossary frame depicted in Figure 1.

An online help screen (Appendix D) was available throughout the reading session, as well as a printed instruction summary (Appendix E). Students were advised to read for comprehension rather than word-for-word translation, and were informed that a written summary post-test would be requested of them after the reading. Subjects were not informed of the vocabulary post-tests. In addition to replicating the conditions of the Chun and Plass (1996) study, this combination of instructions was intended to keep subjects focused on the reading task while avoiding specific attention to intentional vocabulary acquisition. Note that inasmuch as the subjects were not informed about the vocabulary post-test beforehand, this study meets Hulstijn's (2003) methodological definition of incidental learning.

After completing the reading, students were instructed to click the "Exit Story" button. This cleared the reading environment from the screen and opened an online instruction page (Appendix F), explaining that an unannounced vocabulary quiz was about to be administered. Upon clicking the "Proceed" button, subjects were given the *kanji* vocabulary test, followed by the *kana* vocabulary test. After these were completed, a final page was displayed with a writing prompt (Appendix G) for the summary post-test. Upon completion of the summary, students were shown a concluding page thanking them for their participation and asking them not to reveal the existence of the vocabulary post-tests to their classmates.

Data Analysis

Total rates of incidental vocabulary acquisition were determined with simple descriptive statistics. Each vocabulary test answer of each student was treated as an individual data point, and items that were marked with an “Already Knew” response were disregarded. Then, the percentage of correctly answered questions were calculated to determine the overall rate of incidental vocabulary acquisition. Acquisition rates were calculated according to orthography of the words, and according to annotation type, as well as overall rate of acquisition.

Determining the relative effects of annotation type on acquisition rate required more sophisticated inferential statistics. In this analysis, annotation type was the first independent nominal variable at three levels: text-only, text-plus-picture, and text-plus-video. The second independent variable was orthography, which was also nominal and at two levels: *kanji* and *kana*. A third independent variable, with a binomial value of true or false, was included to indicate whether each subject actually viewed the *target* annotation for each keyword.⁴ The dependent variable of vocabulary acquisition was operationalized as the proportion of correct answers on the vocabulary post-tests. Because the response being measured by this study took the form of a proportion, a logistic regression was chosen as the inferential statistical test. Further, because the research design was blocked by subjects, i.e. each subject received all three treatments, and featured two non-continuous independent

⁴ The target annotation is the text annotation for text-only words, picture for text-plus-picture words, or video for text-plus-video words. Note that under this definition, if a subject looked at the text definition of a text-plus-video word, but did not look at the video, the variable for viewing would have a *false* value.

variables, a mixed model was indicated. Thus, the researcher decided upon a mixed models logistic regression.

Using the materials, procedures and statistical tools discussed above, the researcher administered a multimedia glossing treatment to a group of L2 Japanese learners at the university level. Their vocabulary post-test scores were recorded and analyzed, and are reported in the following chapter.

Chapter 4: Results

Research Question 1

The first research question asked: *What is the rate of incidental vocabulary acquisition in Japanese L2 learners after a multimedia glossary treatment?* As described above, acquisition rates were calculated as proportions of correctly answered, previously unknown words from the post-tests. The results for acquisition rate across orthographies and annotation types, along with the number of previously known words, are given in Tables 3 and 4.

Table 3

Incidental Acquisition Rates of Words on First Exposure by Glossary Annotation

	Text	Text+Picture	Text+Video	Overall
Already known	269	236	280	785
Previously unknown	361	394	350	1105
Correctly answered	314	339	316	969
Proportion correct	87.0%	86.0%	89.7%	87.7%

In regard to the data from Table 3, acquisition rates across the board came out slightly higher than those observed by Chun and Plass (1996) in their third study which, like the present study, measured recognition of vocabulary rather than production. Their results indicated an overall acquisition rate of 77.1%, compared to the present study's result of 87.7%. Further, Chun and Plass observed the following rates of acquisition for each type of annotation, in

ascending order: text at 75.1%; text-plus-picture at 76.2%; text-plus-video at 81.4%.

Chun and Plass's findings contrast with the present study's findings in Table 3, which suggest that in this case, text-plus-picture words were recalled least often, while text-plus-video words were recalled most, with text-only words in between. Interestingly, the spread between annotation types was also smaller in this study: the difference between text-plus-picture, which had the lowest rate, and text-plus-video, which had the highest rate, was 3.7%, compared with the high-to-low difference of 6.3% in the Chun and Plass study.

The differences in acquisition rates across annotation types will be treated below in the subsection devoted to Research Questions 2 and 3. However, the differences in overall acquisition rates between the Chun and Plass study and the present one call for some remarks here. The present study was not designed to make causal inferences comparing present acquisition rates with the rates of other studies, as this is not within the scope of the present research questions. Thus, it is not clear why the present study observed a higher overall vocabulary acquisition rate than that observed by Chun and Plass. Several explanations, however, may be conjectured.

Technological-cultural factors are a possible source of subject-related variation: university students in 1996, the year of the Chun and Plass study, may have utilized and related to computers and multimedia differently than students did in 2009, the year of the present study. It is possible that the increasing availability of personal computers, and an attending greater level of technological sophistication on the part of many students, have increased the

effectiveness of multimedia glossing for university students over the past thirteen years.

Another possible source of variation is the language of the text and keywords. There may be factors not measured in this study that make Japanese vocabulary in general more readily learnable than German vocabulary in a multimedia context.

Finally, the treatment may have been a source of variation. While the *CyberHon* application specifications in this study closely matched the specifications of the *CyberBuch* application developed by Chun and Plass, the multimedia content that comprised the glossaries of each was, necessarily, different. It is conceivable that the pictures and video selected for inclusion in the present study's glossary were more effective at promoting vocabulary acquisition than the annotations used in the Chun and Plass study.

Thus, a number of factors could be posited to explain the difference in acquisition rates between the Chun and Plass study and the present one. It must be stressed, however, that future research would be required to establish any causal claims regarding the factors mentioned above.

Table 4

Incidental Acquisition Rates of Words on First Exposure by Orthography

	Kana	Kanji	Overall
Already known	419	366	785
Previously unknown	526	579	1105
Correctly answered	466	503	969
Proportion correct	88.6%	86.9%	87.7%

Returning to the data of the present study, the findings in Table 4, which compares acquisition rates across different orthographic representations, do not have an analog in the Chun and Plass study. Yet they are of relevance to the question of the importance of phonemic form in learning *kanji* characters. Once again, the overall acquisition rate in this study was 87.7%. When examined by orthography type, words written in *kanji* were correctly identified 86.9% of the time, while the *kana* word identification rate was slightly higher, at 88.6%. The closeness of the phonemic and the *kanji* acquisition rates, along with the fact that *kanji* acquisition rates *on the whole* did not exceed phonetic acquisition rates, are consistent with the previously discussed studies and theoretical arguments that indicate an important role of the phonemic component of the characters in *kanji* acquisition.

Research Questions 2 and 3

The second and third research questions asked: *Is there a demonstrable difference between the effects of the various annotation types on vocabulary acquisition rates? If vocabulary acquisition rates are different for the various annotation types, does orthography (hiragana versus kanji, i.e. phonemic versus non-phonemic characters) interact with the main effect of annotation type?* A mixed models logistic regression was conducted to answer these questions. The model is given in Table 5, and its results are shown in Table 6.

Table 5

*Model of Annotation Type and Orthography Effects on Vocabulary Acquisition**

Effect	Orthography	Annotation	Viewed	Standard Estimate	Error
Annotation		video		0	.
Orthography*Annotation	hiragana	picture		1.0110	0.3572
Orthography*Annotation	hiragana	text		0.4002	0.3861
Orthography*Annotation	hiragana	video		0	.
Orthography*Annotation	kanji	picture		0	.
Orthography*Annotation	kanji	text		0	.
Orthography*Annotation	kanji	video		0	.
Viewed			0	-0.6038	0.2214
Viewed			1	0	.
Orthography*Viewed	hiragana		0	0.6118	0.3052
Orthography*Viewed	hiragana		1	0	.
Orthography*Viewed	kanji		0	0	.
Orthography*Viewed	kanji		1	0	.

* Annotation - glossary annotation type (text, text-plus-picture, text-plus-video)
 Orthography - orthographic representation of keyword (hiragana, kanji)
 Viewed - whether or not the target annotation was viewed

Table 6

*Type 3 Tests of Fixed Effects for Annotation and Orthography**

Effect	Num DF	Den DF	F Value	Pr > F
Orthography	1	1848	0.42	0.5172
Annotation	2	1848	10.02	< 0.0001
Orthography*Annotation	2	1848	4.28	0.0139
Viewed	1	1848	3.33	0.0684
Orthography*Viewed	1	1848	4.02	0.0451

* Annotation - glossary annotation type (text, text-plus-picture, text-plus-video)
 Orthography - orthographic representation of keyword (hiragana, kanji)
 Viewed - whether or not the target annotation was viewed

In answer to Research Question 2, the mixed models logistic regression revealed a highly significant effect ($p < 0.0001$) for annotation on vocabulary acquisition. No significant effect for orthography was detected, however a test for interaction was necessary before interpreting the main effects. Interestingly, viewing the target annotation did not have an observable significant effect on vocabulary acquisition, although a near significant result ($p = 0.0684$) was observed.

In answer to Research Question 3, a significant interaction ($p = 0.0139$) was discovered between orthography and annotation type, as well as between orthography and actually viewing the target annotation⁵ ($p = 0.0451$). A plot of the effects is given in Figure 3, indicating the mixed model's predicted proportion of vocabulary items acquired based on orthography and annotation. Two pairs of plot lines are given, both for words whose target annotations were accessed by the subject (henceforth referred to as "viewed") and for those whose target annotations were not accessed by the subject (or "unviewed"). Again, note that while some words are called "unviewed," this means that the *target* annotation was not viewed; other glossary content for that word may still have been accessed.

Interaction between available annotation and orthography. Previous findings (Chun and Plass, 1996; Al-Seghayer 2001; and Iheanacho, 1997) suggested that both picture and video annotations were more helpful than text-only for vocabulary acquisition. In a departure from the observed trend, this study found that for *kanji* words, text-plus-picture words had the lowest recognition rates,

⁵ The present study's use of the term *target annotation* was described in Chapter 3, Footnote 4.

while text-only words had higher rates, and text-plus-video words had the highest rates. As seen in Figure 3, this was observed for both viewed and unviewed *kanji* characters.

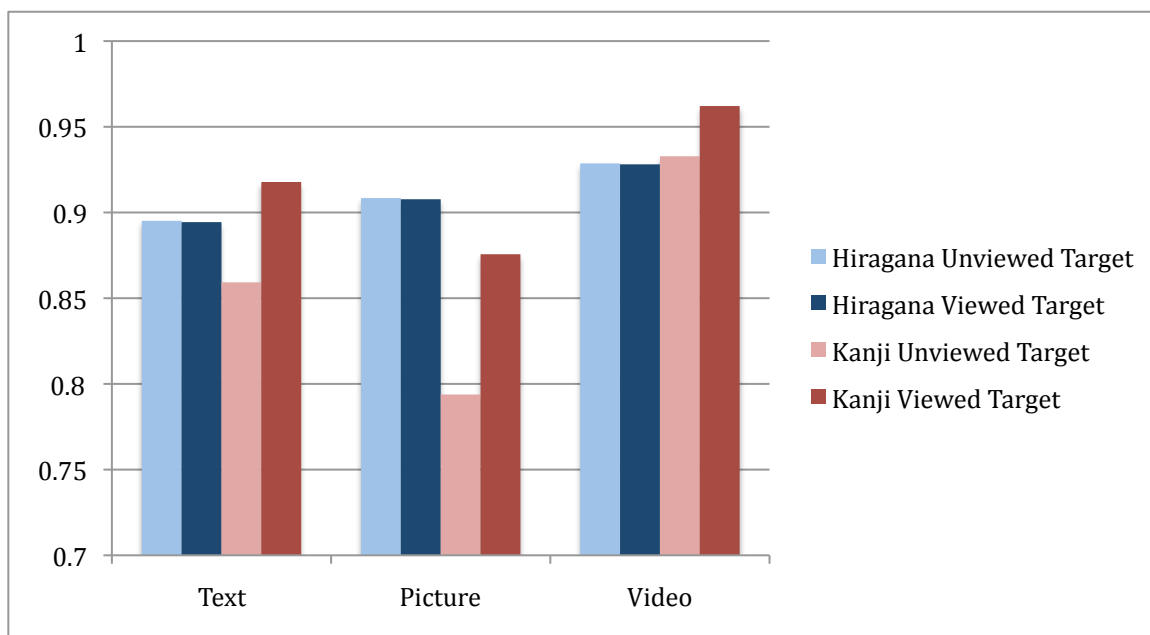


Figure 3. Graph of interaction between orthography and annotation type on proportion of vocabulary items identified correctly.

Hiragana acquisition rates, in contrast to *kanji*, were consistent with previous studies. In the present study, text-only words had the lowest rates of *hiragana* acquisition, while text-plus-picture words had greater acquisition rates, and text-plus-video showed the greatest rates. Again, note that these trends were the same for both viewed and unviewed words.

It thus appears that glossary annotations have different effects on acquisition rate depending on the orthography used to represent the vocabulary item. In other words, the effects of annotation were different for recognition of *hiragana* representations of the keywords than they were for *kanji*

representations. Specifically, *hiragana* acquisition rates were similar to previous studies' results, with multimedia-annotated keywords being recalled more readily than text-only words. *Kanji*, on the other hand, exhibited a previously-unseen pattern, where text-plus-picture annotations were associated with the lowest acquisition rates.

A difference in acquisition patterns could be related to the most obvious difference between *hiragana* and *kanji* – their degrees of phonemic transparency. *Hiragana*, which is similar in phonemic transparency to the alphabetic orthographies of languages from previous studies, exhibited an acquisition pattern similar to the previously observed patterns. *Kanji*, which is unique in its phonemically opacity, exhibited a unique acquisition pattern.

The issue of phonemic transparency may also explain the specific observation of low *kanji* acquisition scores for text-plus-picture annotations. The fact that *kanji* acquisition rates of text-plus-picture words were lower than those of text-only words seems to indicate that the added picture content may have been not only unhelpful, but even detrimental to *kanji* acquisition. Recall that the same was not true for *hiragana* acquisition patterns for the same keywords. This observation reminds us that although the strict classification of *kanji* as ideographs has been discredited by the research and theories discussed in Chapter 2, *kanji* characters and *hiragana* characters still appear to be processed in significantly different ways, at least by L2 learners.

It must be clearly understood that these explanations are speculative and not directly supported by the evidence gathered in the present study. Further, there is no theoretical framework to explain why still pictures would interfere

with *kanji* processing, while video seems to enhance it. Also, as will be discussed below, other data gathered in the present study make any orthographically-rooted explanation of the noted differences problematic.

Returning to the data in Figure 3, it is apparent that *hiragana* acquisition rates across annotation types varied to a much lesser extent than did *kanji* rates across annotations. This may indicate that multimedia annotation affected *hiragana* acquisition less than it affected *kanji* acquisition. Recalling the *kanji-as-ideographs* explanation above, we might suppose that *kanji* processing utilizes visual faculties of the mind and is therefore more susceptible to the effects of visually rich multimedia annotation. The precise mechanisms of such an interaction are beyond the scope of this study, however, and it must be reiterated that this is a purely speculative explanation.

Interaction between available annotation, viewing the target annotation and orthography. Again referring to Figure 3, we see that the interaction between annotation and orthography was different for unviewed words than it was for viewed words. Unviewed words exhibited almost no difference between different orthographies annotated with text-plus-video, a slightly larger difference between words annotated with text-only (0.895 for *hiragana* versus 0.859 for *kanji*), and the greatest difference between words annotated with text-plus-picture (0.908 for *hiragana* versus 0.794 for *kanji*). For both text-only and text-plus-picture words, *hiragana* acquisition rates were the higher of the two.

Viewed words showed a different interaction between annotation and orthography. Text-plus-picture *hiragana* words had higher acquisition rates than text-plus-picture *kanji* words, which was consistent with their unviewed

counterparts, however the difference between the two was much less for viewed words (0.908 for *hiragana* versus 0.876 for *kanji*). For both text-only and text-plus-video words, viewed *kanji* acquisition rates were higher than viewed *hiragana* rates, in contrast with the unviewed acquisition rates in this study. Text-only rates were 0.894 for *hiragana* versus 0.918 for *kanji*, and text-plus-video rates were 0.928 for *hiragana* versus 0.962 for *kanji*. On the whole acquisition rates were more similar across orthographies for viewed words than for unviewed words. This, combined with the overall vocabulary gains from viewing target annotations, seems to support the case that multimedia glossing promotes incidental vocabulary acquisition.

Note that unviewed vocabulary acquisition rates for *hiragana* were higher than or equal to *kanji* acquisition rates for each annotation type. This is consistent with previously discussed studies that suggested phonetic acquisition is a necessary element of *kanji* acquisition. Viewed vocabulary acquisition rates, on the other hand, exhibited a reversal for text-only and text-plus-video words, where *kanji* acquisition rates exceeded *hiragana* acquisition rates. The findings for viewed words thus contradict previous studies that indicated phonemic mastery goes hand-in-hand with the semantic acquisition of *kanji* characters.

This may indicate that, at least in the short-term, glossing in an incidental learning context can in some cases circumvent or overcome the supposedly critical element of phonemic mastery in *kanji* character acquisition. This appears to be true for both text-plus-video and text-only annotations, leading to the conclusion that multimedia content is not necessary for the effect to occur. Two factors unique to the present type of study may be involved in this effect. The

first is the interactive nature of an online glossary, which allows users to access annotations at will, while actively using the keyword in a task. The second factor is learning a keyword in context during an extensive reading task.

Another possible contributing factor must be mentioned in relation to this observation, as well as all previously discussed explanations in this chapter. Subjects may not have been paying attention to the *hiragana* representation displayed in the glossary whenever they clicked a keyword in the text. Since the glossary was intended to be their source of exposure to *hiragana*, they simply may not have allowed themselves to be exposed to *hiragana* for many of the keywords, and thus did not acquire as many *hiragana* representations as they might have.

To justify this explanation, recall the observation that *hiragana* acquisition rates were not as strongly affected by annotation type as *kanji* rates were. Further, observe in Figure 3 that while *kanji* acquisition rates showed a consistent increase between viewed and unviewed instances, *hiragana* acquisition rates appeared almost unchanged whether their target annotations were viewed or not. If the availability and viewing of a keyword's target annotation, whether it was text, picture or video, had little to no effect on *hiragana* acquisition rates, it seems plausible that subjects were simply not paying attention to the *hiragana* representation in the glossary.

A reading task was the vehicle for the treatment in this study, as in previous studies, and subjects were encouraged to read extensively, for comprehension. This may have unintentionally encouraged subjects to ignore the *hiragana* representation displayed when a keyword was clicked in the body of

the text, since learning the pronunciation of the keywords was not strictly necessary for the comprehension task at hand. That is to say, it would have been possible for subjects to click a keyword, view its annotation(s) and insert its meaning into the string of text they were decoding without focusing on the *hiragana* caption in the glossary.

That might partially explain how, contrary to the results of previous studies and theoretical positions, *kanji* acquisition exceeded *hiragana* acquisition in the case of text-only and text-plus-video words where the target annotation was viewed. It might also affect how we consider the explanations previously offered for other observations, since some of the differences in acquisition rates attributed to *hiragana*'s phonemically transparent nature may have been partially caused by inattentiveness of the subjects.

An explanation based solely on subject inattentiveness, however, leaves important questions unanswered. First, if subjects were, by and large, not paying attention to the *hiragana* captions in the glossary, why were *hiragana* acquisition rates so high, at 88.4% after correcting for previous knowledge? Moreover, if subjects did not attend to the *hiragana* captions, how could *hiragana* acquisition rates have surpassed *kanji* acquisition rates, as was the case for text-plus-picture words when the target annotation was viewed? Given these unresolved issues, an explanation based solely on subject inattentiveness seems untenable.

Once again, a complete explanation of the mechanisms underlying the observed differences is beyond the scope of this study. For the present purposes, it suffices to say that there are a constellation of possible effects and explanations

for the observed differences. Ultimately, however, it appears clear that multimedia annotations do have an effect on incidental *kanji* acquisition, and that their effects are not like those observed in previous studies. Further, given the context of the treatment in this study, multimedia annotations do not appear to have as strong an effect on the incidental acquisition of *hiragana* representations of vocabulary.

Chapter 5: Discussion and Conclusions

Incidental Vocabulary Acquisition

Research Question 1 asked about the rate of incidental vocabulary acquisition after subjects received a multimedia glossary treatment. The present study found, first of all, that short-term incidental vocabulary acquisition did occur across all annotation types and orthographies, and that it occurred at an overall rate of 87.7%. That figure is higher than the Chun and Plass (1996) study, of which this was a replication. Several explanations may be posited for this observation, however it is important to note that Chun and Plass's study was conducted under different conditions, with different subjects and slightly different testing methods, thus precluding any direct comparison of vocabulary acquisition rates.

Consequently, the most important conclusion to be drawn in connection to Research Question 1 is that this study provides further evidence that incidental vocabulary acquisition through reading does occur, in this case for Japanese L2 learners. Furthermore, the rate of acquisition, in the receptive form measured here, is quite high; it measured over 85% for each annotation type and for each orthography.

Relative Effectiveness of Multimedia Annotations

Having addressed the overall rate of incidental vocabulary acquisition, let us turn our attention to the effectiveness of multimedia content on vocabulary acquisition. Research Question 2 asked whether multimedia annotations had a demonstrable effect on vocabulary acquisition compared to words annotated

with text only. Although a highly significant main effect was discovered for annotation ($p < 0.0001$), a consideration of further interactions was necessary before drawing conclusions regarding the effectiveness of each annotation type.

This led to Research Question 3, which examined interactions between orthography and annotation. Tests revealed interactions between orthography and available annotation type, as well as between orthography and actually viewing the target annotation. From these results, several interesting observations were made.

First, text-plus-video annotations were associated with the highest vocabulary gains for both *kanji* and *hiragana* representations of the keywords, indicating that video content may be the most effective at promoting incidental vocabulary acquisition. Second, text-plus-picture annotations were associated with lower *kanji* acquisition scores than text-only. This effect was not observed for *hiragana* acquisition, suggesting that picture content, as deployed in this treatment, may have been detrimental to *kanji* acquisition. Third, the fact that annotation type appeared to have a different effect on *kanji* acquisition than it did on *hiragana* acquisition indicates that the two orthographies may be processed in fundamentally different ways, at least by L2 Japanese learners. This suggests the possibility that *kanji* may not be processed in a strictly phonemic sense. Fourth, it appears that despite previous evidence demonstrating the importance of phonemic mastery in *kanji* acquisition, semantic acquisition of a *kanji* character does, in some cases, occur without phonemic acquisition.

Some of these observations ran counter to the expectations derived from the review of the current literature on the subjects of multimedia glossing and

kanji acquisition. While several possible explanations for these unexpected observations were offered, the present study was not designed to explain the mechanisms that caused the observed results. Future research will be required to confirm the present results and to test for causal relationships not addressed in this research project.

Implications

Despite some unanswered questions, certain implications seem clearly borne out by the facts at hand in this study. First, it appears that, consistent with Krashen's theory and other previous research, incidental acquisition of *kanji* and *hiragana* representations of words does occur through extensive reading when supplemented by glossing. Further, it occurs at relatively high rates, as measured by short-term recognition tests. This suggests that extensive reading tasks may be an effective means of "teaching" vocabulary in third- and fourth-year university curricula where receptive mastery is valued.

Japanese literature teachers, in particular, may wish to consider utilizing glossary-enabled texts where possible. Previous glossing research and the present study both point toward the benefits of glossing, and multimedia glossing in particular, for incidental vocabulary acquisition. We may conclude that a multimedia glossary-enriched reading curriculum could, with no additional assignments to burden the teacher or student, increase vocabulary acquisition on the part of learners.

Second, consistent with previous studies, it appears that different kinds of multimedia glossary content have different effects on *kanji* acquisition. When one comes to the specifics of these effects, however, the results of the present

study diverge somewhat from previous studies. Based on the present results, text-plus-video annotations seem to facilitate greater *kanji* acquisition than text-only annotations, while text-plus-picture may have a deleterious effect.

Further investigation is required to clarify this connection. For example, future studies could examine the characteristics of specific picture and video annotations to see how those characteristics affect vocabulary acquisition. Additionally, future studies could examine how the characteristics of the keywords interact with multimedia content to affect vocabulary acquisition.

If the trends observed in this study are consistently exhibited in future research, designers of Japanese instructional materials and instructors of Japanese as a foreign language may wish to consider incorporating more video-based glossary aids into their reading tasks, while avoiding picture-based aids. Note that although *hiragana* acquisition rates were similar to previously observed patterns, the questions raised above, combined with the minuscule observed effects, tend to discourage any prescription for media content in regard to *hiragana* acquisition.

Third, although previous research and theory indicate that the *kanji* characters are inextricably tied to their phonemic values, both in native readers and in foreign learners, the results of this study indicate that *kanji* may not be processed in the same way as the phonemic *hiragana* representations of words, at least by L2 learners during an extensive reading task involving a glossary. So while the classification of *kanji* as ideographs has been discredited, the results of this study indicate that there is still a significant difference in the way L2 learners acquire *hiragana* and *kanji*.

Perhaps the best classification scheme for orthographies like the *kanji* and the *kana* would be along a continuum. On one end would be a hypothetical, perfectly phonemic orthography, with only phonemic and no visual referents to meaning. On the opposite end of the continuum would be a hypothetical, perfectly ideographic orthography, with a purely visual pointer toward meaning, and no phonemic intervention. Instead of falling neatly into one category or the other, orthographies may have a tendency toward either end of the continuum. In such a scheme, *kana* characters would fall closer to the phonemic end of the scale, while *kanji* might be placed closer to the ideographic end. This model would help explain the differences observed between the two, while still allowing them to share some degree of commonality.

Fourth, it would appear that the nature of a learning task and the materials involved may impact the role of phonemic mastery in *kanji* acquisition. The results of the present study appear to indicate that, contrary to the results of previously conducted studies, semantic mastery of a *kanji* character can occur without phonemic mastery, at least under certain conditions. Specifically, this result was observed when the learner viewed text-only and text-plus-video annotations.

It is possible that the *kanji* acquired without phonemic mastery reflected gains over and above that which would have occurred without the *CyberHon* treatment. Alternatively, it is possible that the *CyberHon* treatment was detrimental to the phonemic mastery of those *kanji*. If the latter possibility proves true, curriculum designers must take care to weigh this potential decrease in phonemic acquisition against the benefit of overall vocabulary gains when

deciding whether to utilize multimedia glossary materials. Some situations may call for a combination learning treatment that incorporates other types of learning activities that would promote the phonemic mastery of key vocabulary items.

Limitations

Regarding Research Questions 2 and 3, one methodological limitation that may have affected this study's outcome was the subjective nature of some elements of the stratification procedure that determined the keywords for each annotation group. Particularly problematic was the rating of words according to difficulty and importance to the sentence in which they were found. The fact that *kanji* acquisition rates appeared to be affected by annotation type, even when the target annotation was not viewed, may indicate that the perceived effect could at least partially have been the result of insufficient homogenization of the annotation groups.

Another factor that may have affected group homogenization was word frequency. While every attempt was made to keep frequencies similar across groups, it should be noted that a perfect balance was not possible given the text, number of keywords and the criteria for choosing keywords. This may have affected lookup behavior and exposure to *kanji* and *hiragana* instances of the keywords, which could have in turn affected acquisition. Future studies might be designed to improve homogenization in terms of frequency.

It should be noted, however, that if the *kanji* acquisition patterns were simply the result of some groups being comprised of "easier" words than others, one would expect *hiragana* acquisition rates to vary in a similar pattern. This was

not the case, as *hiragana* rates were dissimilar to *kanji* in both spread across annotation types and order of effectiveness. Thus, while the homogeneity (or lack thereof) of the annotation groups may have influenced the results, it seems unlikely that this accounts for the entirety of the effect.

A second limitation of the present study involved the process of choosing keywords, which resulted in a bias toward nouns (see Table 2). Recall from the discussion in Chapter 3 that an equal number of words for each annotation category was sought after, to create three homogenous groups of keywords. Also discussed in Chapter 3 was the choice of *Naruhodo* for the present treatment. Of the stories available for inclusion in the study, this one was deemed most appropriate to the reading ability of the subject pool, and it happened to contain a body of viable keywords that was comprised of mostly nouns. Future research using different stories and different types of keywords could be conducted to address this limitation.

Another limitation of this study's methodology involved the logistics of administering the vocabulary post-tests. In many of the previous studies, researchers conducted both immediate and delayed post-tests to measure not only short-term vocabulary acquisition but also long term retention of the acquired words. The present study originally included plans for a delayed post-test, but unexpected complications arose that precluded that possibility. Thus, no statements regarding long-term effects of multimedia glossing could be made.

The following limitations are not the result of methodological issues, but rather the result of the present study's focus. As discussed above, the issues of user behavior and attention, as well as orthographic phonemic transparency,

may be important keys to answering some unanticipated questions that arose during the analysis of the study's data. The present study, however, was not designed to address these questions, and future research will be required to clarify those points.

Also, it must be noted that while the present study focused on psycholinguistic aspects of multimedia glossary effects – how learners mentally process and recall new vocabulary items – the present results do not necessarily reflect solely psycholinguistic effects. Multimedia glossary content may have had other effects on the learning process, such as increasing or decreasing motivation to acquire vocabulary, or affecting the subjects' attentiveness. The present study, like the line of research upon which it was based, was not designed to investigate these affective elements of multimedia glossing. Future research could be conducted to try to separate out the psycholinguistic effects from the affective and other potential effects on vocabulary acquisition.

Recommendations for Future Research

Future research on this subject could eliminate some methodological weaknesses of the present study. A delayed post-test to strengthen the validity of vocabulary acquisition claims would be the easiest to implement. Internal validity would also be strengthened by finding or devising more objective procedures that could determine both vocabulary item difficulty and the importance of a given word in a sentence for the Japanese language, which would help ensure a more homogenous grouping of words for each annotation. A closer study of picture and video appropriateness, or how to choose optimal multimedia content, could also strengthen validity. Finally, finding a story that

provides keywords with a more diverse distribution of parts of speech would improve generalizability.

Subsequent research involving multimedia glossing among Japanese language learners should also address user behavior issues that were not included in the planned scope of this study. Specifically, if future researchers could conduct a similar study, but implement measures to determine whether subjects are actually looking at the *hiragana* representations offered for each glossary entry, they might be able to begin answering the question of why *hiragana* acquisition was unaffected by multimedia annotations in this study.

Finally, future research might aim to discover affective and other effects of multimedia glossing rather than focusing solely on psycholinguistic effects. If non-psycholinguistic effects account for a significant portion of the benefits associated with multimedia glossing, other treatments or teaching methods could be devised to create the same benefits associated with multimedia glossing in other, possibly non-multimedia contexts.

References

- Abraham, L. B. (2001). The effects of multimedia on second language vocabulary learning and reading comprehension. *Dissertation Abstracts International, A: The Humanities and Social Sciences*, 62 (2), 535-A. (UMI No. DA3003439)
- Al-Seghayer, K. (2001). The effect of multimedia annotation modes on L2 vocabulary acquisition: A comparative study. *Language Learning & Technology*, 5(1), 202-232.
- Aoyama, K. (2005). *Using a diglot reader to teach kanji: The effects of audio and romaji on the acquisition of kanji vocabulary*. Unpublished doctoral dissertation, Brigham Young University, Provo, UT.
- Chun, D. M., & Payne, J. S. (2004). What makes students click: Working memory and look-up behavior. *System*, 32(4), 481-503.
- Chun, D. M., & Plass, J. L. (1996). Effects of multimedia annotations on vocabulary acquisition. *Modern Language Journal*, 80(2), 183-198.
- DeFrancis, J. (1984). *The Chinese language: Fact and fantasy*. Honolulu, HI: The University of Hawaii Press.
- Ekstrom, R.B., French, J.W., Harman, H.H., & Dermen, D. (1976). *Manual for kit of factor-referenced cognitive tests*. Princeton, NJ: Educational Testing Service.
- Erbaugh, M.S. (Ed.). (2002). *Difficult characters*. Ann Arbor, MI: The Ohio State University National East Asian Language Resource Center.

- Everson, M. (1998). Word recognition among learners of Chinese as a foreign language: Investigating the relationship between naming and knowing. *Modern Language Journal*, 82(2), 194-204.
- Everson, M. & Ke, C. (1997). An inquiry into the reading strategies of intermediate and advanced learners of Chinese as a foreign language. *Journal of the Chinese Language Teachers Association*, 32(1), 1-20.
- Hulstijn, J. H. (2003). Incidental and intentional learning. In C.J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 349-381). Malden, MA: Blackwell Publishing.
- Hulstijn, J. H., Hollander, M. & Greidanus T. (1996). Incidental vocabulary learning by advanced foreign language students: The influence of marginal glosses, dictionary use, and recurrence of unknown words. *Modern Language Journal*, 80(3), 327-339.
- Iheanacho, C. C. (1997). Effects of two multimedia computer-assisted language learning programs on vocabulary acquisition of intermediate level ESL students. Dissertation Abstracts International, Section A: The Humanities and Social Sciences, 62(8) (UMI No. 3024746)
- Knight, S. (1994). Dictionary use while reading: The effects on comprehension and vocabulary acquisition for students of different verbal abilities. *Modern Language Journal*, 78(3), 285-299.
- Koda, K. (2005). *Insights into second language reading*. Cambridge, UK: Cambridge University Press.

- Krashen, S. D. (1983). *The natural approach: Language acquisition in the classroom*. New York: Pergamon Press.
- Krashen, S. D. (1993). *The Power of reading: Insights from research*. Englewood, CO: Libraries Unlimited.
- Lyman-Hager, M., Davis, N., Burnett, J. & Chennault, R. (1993). Un vie de boy: Interactive reading in French. In F. L. Borchardt & E. M. T. Johnson (Eds.), *Proceedings of the CALICO 1993 annual symposium on "assessment"* (pp. 93-97). Durham, NC: Duke University.
- Matsunaga, S. (1995). *Role of phonological coding in reading kanji: A research report and some pedagogical implications (no. 6)*. Honolulu: University of Hawaii at Manoa.
- Mori, Y. (1998). Effects of First Language and Phonological Accessibility on Kanji Recognition. *The Modern Language Journal*, 82(1), 69-82.
- Nagy, W. E., Herman, P.A. and Anderson, R. (1985). Learning words from context. *Reading Research Quarterly*, 20(2), 233-253.
- Paivio, A. (1971). *Imagery and verbal processes*. New York: Holt, Rinehart and Winston.
- Unger, J.M. (2004). *Ideogram: Chinese characters and the myth of disembodied meaning*. Honolulu: University of Hawai'i Press.
- Yoshii, M., & Flaitz, J. (2002). Second language incidental vocabulary retention: The effect of picture and annotation types. *CALICO Journal*, 20(1), 33-58.

Appendix A: Informed Consent Document

Please note that by accessing the CyberHon website and submitting your results, you indicate that you accept the terms and conditions below:

Consent to be a Research Subject

This study is being conducted by Brian James under the supervision of his faculty mentor, Dr. Paul Warnick, professor of Japanese in the Brigham Young University Asian and Near Eastern Languages Department. The research is in connection with a masters thesis at BYU to pilot test a new multimedia-enabled electronic text. You are being invited to participate because you are a returned missionary from a Japanese-speaking mission enrolled in a third-year Japanese reading course.

You will be asked to read a short story using the multimedia electronic text, and complete a series of post-reading exercises. The reading exercise, survey and test will be administered at JFSB B161 from February 27 to March 6, 2009, and will take approximately 60 to 90 minutes to complete.

Results of the exercises will remain completely confidential: No names will be referred to at any point in the published results of the research project, and results linked with personally identifying information will be stored on private servers administered by the BYU Center for Language Studies. Participants in the study will be referred to by an arbitrarily assigned number in both statistical analysis and any published reports, and any information that links personally identifiable information to test scores will be deleted at the completion of the project (projected April 2009). Results may be retained in an anonymous aggregate format for future research.

Our research will help us to determine the effectiveness and worth of certain kinds of computer-based language learning activities for Japanese learners. We hope this information will help us and other Japanese teachers formulate more effective methods of teaching the language to nonnative learners. Furthermore, participants in this study may receive extra credit for their participation.

Participation in this research study is voluntary. You have the right to withdraw at anytime or refuse to participate entirely without jeopardy to your class status, grade or standing with the university.

If you have questions regarding this study, you may contact Brian James at [REDACTED]. If you have questions you do not feel comfortable asking the researcher, you may contact Christopher Dromey, PhD, IRB Chair, 422-6461; 133 TLRB, Brigham Young University, Provo, UT 84602 USA; [REDACTED].

Appendix B: Questionnaire

Name: _____

Student ID Number: _____

Gender: _____

Age: _____

Have you served an LDS mission in Japan? _____

What Japanese classes are you currently enrolled in? _____

What Japanese classes have you previously taken at the university level?

Appendix C: Text of Login Screen

Thank you for participating in the CyberHon research project. Using this online reading application, you will be asked to read "Naruhodo," a short story by Hoshi Shinichi. The CyberHon application is equipped with a side-bar multimedia-capable glossary, which is linked to a number of keywords (highlighted in blue text) throughout the story. To view the glossary entry for any of the keywords, simply click the word in the text that you wish to view, then select the desired glossary content (text, audio, pictures or video) from the buttons above.

As you read, you will notice that not every annotation type is available for every word, however we ask you to use as much of the glossary and its content as possible. After the reading, you will be asked to complete a short reading comprehension exercise in the form of a summary, to be written in English, as well as a survey soliciting your feedback about your experience using CyberHon. If you have questions at any point during the reading, please click the "Help" button at the bottom of the screen, or ask the test center proctor. Again, thank you for your participation and enjoy the story!

Appendix D: Online Help Screen

Annotation Buttons

Clicking on highlighted terms in the text will activate buttons for available annotations, including text, audio, images, and video.

Glossary Frame

Vocabulary words clicked in the text frame will appear in the glossary frame both in Kanji and Hiragana. Buttons for available annotations will be activated at the top of the application.




Help/Exit Buttons

Text Frame

Use the next and previous buttons to page through the text of the story. Certain vocabulary words are highlighted. Clicking on these words will bring them up in the glossary frame and activate buttons for available annotations.

Appendix E: Instruction Sheet

Student Instructions

1. Log into the Macintosh Operating System environment with your BYU Net ID and password.
2. Open the **Safari** web browser application (please do not use another browser).
3. Go to the following webpage: 
4. Enter your personal information in the fields to the right of the page and click “Begin.”
**** Enter your student ID without any spaces or dashes – just the 9 digit number!**
5. Remember the following points as you read the short story:
 - a. Read for overall story comprehension rather than word-for-word translation.
Try to finish the story within 45 minutes.
 - b. Use the hyperlinked multimedia glossary to look up words with which you are unfamiliar.
 - c. Please refrain from using any other dictionaries or similar tools while reading – do the best you can with the tools provided on the webpage.
6. When you finish reading, click the “Exit Story” button. You will be guided through a series of post-reading exercises.

Thank you for your participation in this study!

Appendix F: Unannounced Vocabulary Quiz Instructions

Before you begin the summary-writing exercise, we would like you to complete a brief vocabulary quiz. The nature of our experiment prevented us from telling you about this quiz before the reading session, but we ask that you please answer the following questions to the best of your ability. You will be asked to provide the English equivalent of a series of words from the story, as well as which glossary content first came to mind for each item on the quiz. Also, we ask for your cooperation in not discussing this 'surprise' quiz with any of your classmates, as it might adversely affect the results of the study. Thank you very much for your cooperation.

Appendix G: Written Summary Prompt

Please write a brief summary, in English, of the story you have just read. Include the main points of the plot as you understood them, but do not worry about capturing minor details. Try to limit your summary to no more than three paragraphs.