This is to certify that the
thesis entitled

EFFECTS OF L2 PROFICIENCY ON WORD RECOGNITION: THE USE OF PHONOLOGY IN READING BY L1 CHINESE LEARNERS OF ENGLISH

presented by

RYAN T. MILLER

has been accepted towards fulfillment
of the requirements for the

MA. degree in Teaching English to Speakers of Other Languages

Diana Pulido
Major Professor's Signature

6-15-2009
Date

PLACE IN RETURN BOX to remove this checkout from your record.
TO AVOID FINES return on or before date due.
MAY BE RECALLED with earlier due date if requested.

DATE DUE	DATE DUE	DATE DUE

EFFECTS OF L2 PROFICIENCY ON WORD RECOGNITION: THE USE OF PHONOLOGY IN READING BY L1 CHINESE LEARNERS OF ENGLISH

By

Ryan T. Miller

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

MASTER OF ARTS

Teaching English to Speakers of Other Languages

2009

ABSTRACT

EFFECTS OF L2 PROFICIENCY ON WORD RECOGNITION: THE USE OF PHONOLOGY IN READING BY L1 CHINESE LEARNERS OF ENGLISH

By

Ryan T. Miller

Much of the way words are recognized and processed in any language is related to and influenced by the writing system used in that language. A multitude of research has shown that first language orthography can have a significant impact on the way that words are recognized in a second language. However, is this transfer effect from the L1 orthography a static one that is insurmountable, or is it possible for cognitive processes to be restructured for the benefit of the L2 reader? The present study investigated this question through the use of two independent measures of word recognition processes. A lexical decision task incorporating an alternating-case text paradigm, and including word-based independent variables of word frequency and spelling regularity indicated that at higher proficiency levels, L1 Chinese learners of English have improved intra-word processing, especially among words with regular grapheme-phoneme correspondence. In addition, a semantic category judgment task using a homophone paradigm indicated that high proficiency readers have a greater reliance on phonology, and a lesser reliance on orthography, than do low proficiency readers.

TABLE OF CONTENTS

LIST OF TABLES....vi

LIST OF FIGURES....vii

Introduction 1

Review of Research 3
- Introduction 3
- English Orthography 4
- Chinese Orthography 5
- Models of Word Recognition 8
  - Visual Route 9
  - Phonological Route 11
  - Dual Route 15
  - Orthographic Depth Hypothesis 16
- Transfer 17
- Proficiency 20

Research Questions 24

Hypotheses 25

Methodology 26
- General Procedure 26
- Background Questionnaire 28
- Proficiency Measure 29
- Participants 30

Experiment 1 35
- Materials 35
- Procedure 39
- Results 40
  - Reaction Times 41
- Accuracy 48
- Discussion 50

Experiment 2 54
- Participants 55
- Materials 55
- Procedure 57
- Results 58

Accuracy 59
Reaction Times 62
Discussion 65

General Discussion 69
Limitations and Future Research 71

Appendices 74
Appendix A 74
Appendix B 75
Appendix C 77
Appendix D 78
Appendix E 80
Appendix F 81
Appendix G 83

References 84

LIST OF TABLES

Table 1 *Means (and standard deviations) of proficiency measure results* 34

Table 2 *Means (and standard deviations) of participant age and length of residence in English-speaking locations* 35

Table 3 *Descriptive statistics for frequency, length, and word class for each condition* .39

Table 4 *Means (and standard deviations) of reaction times and error rates* 42

Table 5 *Summary of means, standard deviations, and coefficients of variation for alternating case stimuli* 45

Table 6 *Examples of phonological and orthographic similarity* 57

Table 7 *Means and standard deviations of error rates for each condition, in percents* ...60

LIST OF FIGURES

Figure 1. Chinese and English orthography. 5

Figure 2. Direct or visual route to meaning (post-lexical phonology). 9

Figure 3. Phonology-mediated route to meaning (pre-lexical phonology). 12

Figure 4. Dual-route model. 16

Figure 5. MTELP scores for all 34 participants (maximum possible score of 60). 31

Figure 6. Interaction between spelling regularity and proficiency within case. 44

Figure 7. Means of lower case and alternating case reaction times by proficiency group..46

Figure 8. Interaction between case and proficiency within spelling regularity conditions...47

Figure 9. Error rate interaction between proficiency group and frequency. 49

Figure 10. Paradigm used in Experiment 2. 58

Figure 11. Category judgment accuracies in the four conditions by proficiency group. .61

Figure 12. Error rates of phonological similarity in the similarly-spelled and less similarly spelled conditions. 64

Figure 13. Mean reaction times for phonological similarity by proficiency group. 65

Introduction

Reading, in any language, is a complicated process that involves a wide range of subprocesses in order to be successful, and a skill that has long been of interest to researchers. These subprocesses of reading are typically divided into so-called higher-level processes and lower-level processes. Higher-level processes involve such concepts as how ideas within the text relate to each other (cohesion), the organization of the parts of a text (rhetorical structure), and what the main ideas of a text are and how they relate to the reader's own knowledge of the world. Lower-level processes involve the processing of the pieces that make up a text, such as letters, words, characters, and the smaller units that these are comprised of. Which type of processing is more important is a question that has long been debated. Top-down models (that is, models that investigate how higher-level processes affect lower-level processes) have largely focused on the importance of background knowledge of content and patterns of structure, while bottom-up models have focused on automaticity of word recognition, or the process of understanding meaning from the graphemic representation of a word.

While strong views of top-down and bottom-up processing, where one type of processing feeds into the other in a serial manner, were once prevalent, it is now widely accepted that there is interplay between the two. Stanovich (1981) proposed an interactive-compensatory model in which deficiencies in either higher- or lower-level processing can be made up for by the other. This can be exemplified by the concept of hypothesis-testing. Here, the reader makes a hypothesis about what is to come based on background knowledge; when there is less background knowledge available for top-down processing, the reader is less able to make an accurate hypothesis. In this case, the reader is required to gain as much information as possible from the text itself. On the other hand, deficient lower-level processing can be facilitated by greater background knowledge that would allow the reader to make a better hypothesis. However, this greater use of background knowledge to compensate for inadequate lower-level processing consumes attentional resources (Stanovich, 1980), and because the human brain has a limited amount of cognitive capacity (LaBerge & Samuels, 1974; Schnieder & Shiffrin, 1977), the greater consumption of cognitive resources to form hypotheses can mean that fewer resources are available for comprehension (Stanovich, 1980). As such, deficient lower-level processes can greatly hinder comprehension in a way that it is difficult for higher-level processing to compensate for, and can even make comprehension impossible (Segalowitz, 1986). It is for this reason that word recognition is often seen as the foundation of reading comprehension (Gough, 1984; Stanovich, 1991).

For second language readers, the word recognition process can be further complicated by differences inherent in the processing of their first and second languages. It has been suggested that different orthographies entail differences in word recognition, and that word recognition strategies are formed in a reader based on the orthography of their first language (Durgunoglu, Nagy, & Hancin-Bhatt, 1993). First language word recognition processes have been shown to influence second language learner reading, especially when that second language uses an orthography that is typologically different from that of the reader's first language (Koda, 1988, 1989, 1990).

This influence from the learner's L1 has been considered to be a static element in the previous research, with consideration being given almost exclusively to the characteristics of the L1 and the L2. While this research has been fruitful in describing the nature of word recognition in different languages and the influence of one language upon another, there has been little attention paid to the role that the readers themselves play, and how individual differences in the reader may affect word recognition. Recently, however, more research has taken note of the possibility that word recognition processes may actually be more fluid than previously thought, and with increased exposure to and greater proficiency in a second language, the reader's word recognition and processing of that language might change (e.g. Akamatsu, 2005; Chikamatsu, 2006). While much of the previous research has focused on *how* the L1 affects L2 word recognition, in order to have a more comprehensive understanding of L2 word recognition processes, we must also investigate *to what extent* L1 orthography influences L2 word recognition in individual readers. Koda (2005) emphasizes this point by stating that there is a need for studies to investigate "developmental changes in the **manner** and **extent to which** L1 and L2 processing experiences coalesce in forming L2 word-recognition skills" (p.46; emphasis added). The present study seeks to explore this important, but as of yet largely neglected, aspect of reading by investigating the effect of proficiency on the use of phonology during word recognition by L1 Chinese learners of English.

Review of Research

Introduction

Word recognition processes in any written language have been shown to be shaped by the nature of the writing system of the language. At the most basic level, the various writing systems of the world's languages map spoken language to written characters in three main ways (DeFrancis, 1989; Muljani, Koda, & Moates, 1998; Wang, Koda, & Perfetti, 2003): alphabetic orthographies map phonemes, the smallest units of sound; syllabic orthographies map syllables, or units of multiple sounds; and logographic/morphographic orthographies map words or morphemes, the smallest unit of meaning (Spencer, 1991). These fundamental differences in the orthographies of languages have significant implications for the way that words are recognized in each language.

English Orthography

In English and other alphabetic languages, words consist of letters, which each have a sound that they represent. In these languages, there are strong relationships between letters and their sounds, called grapheme-phoneme correspondence (GPC) rules. Because English words can be processed serially, letter by letter, GPC rules make the phonological information of each constituent letter of a word accessible during decoding. When reading English words, native English speakers assemble the sounds of the letters to form words; this is how native-English speaking children learn to read (Ehri, 1998), and is the fundamental process of word recognition in English.

However, even though the phonology of words is highly accessible in English, very familiar words (also called sight-words) have been shown to be recognized based on their visual representation, without need for activation of phonology. In fact, research has shown that highly-skilled readers could name familiar, monosyllabic words (made up of multiple letters) just as quickly as they could name a single letter (Doehring, 1976). In addition, relying on GPC rules alone does not always give adequate information, as in the case of irregularly spelled words such as ‘knight’ or ‘colonel’.

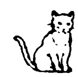
	Chinese	English
Orthographic representation	猫	c a t
Phonological representation	/ mao/ 1	/k/ / æ/ /t/
Meaning		

Figure 1. Chinese and English orthography.

Chinese Orthography

The Chinese language has long been considered to be a logography. Although English words have grapheme-phoneme correspondence rules that the reader can apply to pronounce and assemble each letter of a word, this concept does not apply to Chinese, as words are formed in a single block of a number of strokes that do not map out a full phonetic representation of the word (see Figure 1).

At a phonological level, rather than a written character representing a single phoneme, each Chinese character represents a syllable. In addition, at a semantic level, each character represents a morpheme, or a unit of meaning. This is different from English in that, as a constituent letter of a word, the letter 'f' has no meaning other than to indicate a labio-dental fricative in English, whereas a single Chinese character can represent an entire concept, e.g. 火 (/huo/3[1]), or *fire*. Such a character cannot be broken down into any smaller unit. Chinese words are generally considered to be

[1] Throughout this paper, numbers will be used to represent the tones of standard Mandarin Chinese. The numbers 1, 2, 3, and 4 will be used to represent the level (平聲), rising (上聲), departing (去聲), and entering (入聲) tones, respectively.

monomorphemic, although polymorphemic (and thus multi-syllabic and multi-character) words also exist. It is because of this relationship between the written form, sound, and meaning that Chinese characters are often referred to as morphemic (e.g., Leong, 1986) or morphosyllabic (e.g., DeFrancis, 1989).

In addition, modern standard Chinese consists of 4,574 characters (Language and Teaching Institute of Beijing Linguistic College, 1986, as cited in Perfetti, Liu, & Tan, 2005). Also, in modern standard Chinese, there are roughly 420 different syllables (disregarding tones), and thus for every syllable there are, on average, about eleven characters that share the same syllable (i.e. homophones). If the Chinese writing system had been a representation of the phonology of the language, the reader would then need to disambiguate between these eleven homophones using syntactic or semantic context, thus consuming more cognitive resources (Katz & Frost, 1992). Such a system would be far from efficient. This would seem to indicate that, during reading, greater reliance on orthographic representation would be necessary as reliance on phonological representation would only lead to confusion among the ten other words that have the same pronunciation (Katz & Frost, 1992). Since these homophones are not homographs, however, reliance on the orthographic representation would give sure-fire information about the intended meaning.

On the other hand, if Chinese characters had no systematic relation to phonology, a reader would have to remember, without any cue, the phonology of each and every one of the 4,574 characters individually, which would be a very daunting task. Although many characters in Chinese consist of only a single part (such as 'fire' above), a large number of characters are actually an amalgam of smaller pieces put together; such characters are called compound characters, and the smaller pieces are called radicals. Most compound characters consist of at least two radicals, but can contain many more. Radicals can be divided into two types: semantic radicals and phonetic radicals. Semantic radicals provide a hint as to the meaning of the compound character, while phonetic radicals provide a hint as to the pronunciation. For example, 日 is pronounced /ri/4, and has the meaning of *sun*. The character 青, pronounced /qing/1, has the meaning of *green*. A combination of 日 (/ri/4, *sun*) on the left and 青 (/qing/1, *green*) on the right produce the character 晴, pronounced /qing/2 and meaning *sunshine*. In this character, both a semantic radical (sun, indicating a semantic relationship between the radical [sun] and the compound [sunshine]) on the left and a phonetic radical (/qing/1) on the right (indicating a phonetic relationship between the radical [pronounced /qing/1] and the compound [pronounced /qing/2]) are present (Perfetti, Liu, Fiez, Nelson, Bolger & Tan, 2007). Overall, there are 895 phonetic radicals and 214 semantic radicals in Chinese (DeFrancis, 1989). Unlike the previous example, however, most compound characters have only one radical that is either semantic or phonetic, but not both. The rest of the character typically has no relation, either semantic or phonetic, to the character as a whole. Although the greater number of phonetic radicals may seem to indicate that reliable phonetic cues are embedded in Chinese characters, phonetic radicals are reliable only 38% of the time (Zhou, 1978, as cited in Perfetti, Liu, & Tan, 2005). In addition, there are many compound characters that contain neither a phonetic nor a semantic radical. For example, 青 (/qing/1, green) is actually a compound character that consists of two characters that provide neither a semantic nor a phonetic cue: 生, meaning *life* and pronounced /sheng/1, on the top and 月, meaning *moon* and pronounced /yue/4, on the bottom. Thus, although both semantic information and phonological information can be indicated by the radicals within compound characters, this information is often unreliable, and even when it is reliable, understanding it requires parsing that is much different from the way that English words are parsed.

Models of Word Recognition

Before it is possible to investigate how writing systems affect word recognition processes of L2 learners, there must first be a description of the available models of word recognition. Psycholinguistic models of word recognition will be discussed in terms of the major features, strengths, weaknesses, and assumptions that underlie them, and supporting research will be presented for each.

All models of word recognition incorporate, at some point, three basic elements: orthography, phonology, and semantics (DeFrancis, 1989; Wang, Koda, & Perfetti, 2003), and according to the Universal Phonological Principle (Perfetti, 2003; Perfetti, Zhang, & Berent, 1992), there is a direct relationship between orthography and phonological activation in the accessing of semantic information in all languages. What has been under debate for a number of years is the order in which, and the degree to which, these are processed or activated. Three main types of psycholinguistic models of phonological involvement in word recognition have emerged: strong visual-route models, strong phonological-route models, and dual-route models. Although these models will be discussed here mostly in terms of first language word recognition, they also have important implications for L2 reading in the sense that transfer of L1 word recognition processes has been shown to occur in second language readers. The issue of transfer will be discussed later.

Visual Route

In the strong visual-route model of word recognition, the meaning of a word is accessed directly from the orthography. That is, a written word is viewed and the overall visual shape of the word is recognized, which then leads to access of the meaning of the word, that is, lexical access (Figure 2). Unlike phonological models, where phonology is activated between recognition of the orthography and semantic access, phonological activation has been proposed by various visual-route models to occur after lexical access, at the same time as lexical access, or not at all. These are also referred to as direct-route models because semantic knowledge is accessed directly from the orthography, without pre-lexical activation of phonology; this type of recognition is also referred to as addressed recognition.

The biggest obvious advantage of the visual-route model is that the intermediary step of accessing phonology that phonological routes propose is cut out of the process. This streamlined method would seem to be more efficient than a phonology-mediated route of access, and, in fact, some studies (e.g. Brown and Haynes, 1985; Fender, 2003) have suggested that accessing the lexicon through a direct route does result in faster word

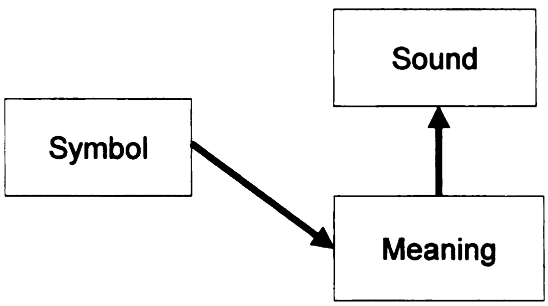


Figure 2. Direct or visual route to meaning (post-lexical phonology).

recognition.

Perhaps the strongest support for visual-route models is found in studies of the effects of word frequency on word recognition. This is because, according to a strong phonological-route model, words are assembled piece by piece, and therefore it should not matter if an entire word is more frequently encountered or not; if anything, word length or frequency of certain letter combinations should be more important. Baluch and Besner (1991) conducted a study that supports a visual route of word recognition based on word frequency effects, investigating word recognition in Persian. Persian was used for this study because some words include vowel sounds that are specified in the orthography, whereas other words include vowel sounds that are not, thus creating a variation in the degree of orthographic transparency within the single script. Baluch and Besner found that there were effects of word frequency not only in the words that were more orthographically opaque, but also the orthographically more transparent words as well. This evidence was interpreted as showing that even when a script does allow a one-to-one matching with the phonology (and thus would be fairly easy for the reader to parse letter by letter), words are still recognized by their overall form, thus resulting in effects of frequency. Similarly, research has shown that phonology plays a lesser role in stronger readers than weaker readers (through the existence of so-called 'sight words') (Chen, Flores d'Arcais, & Cheung, 1995; Ehri, 1998) .

If a phonological route were to explain the existence of familiarity effects, it would need to show that it is the frequency of letter combinations matching with sounds that would assist in word recognition, rather than the frequency of the entire word (as, in a strong phonological-route model, the word would be parsed serially). On the contrary, Peereman, Content, and Bonin (1998) found that, for French speakers (a fairly transparent orthography), spelling regularity (that is, the consistency of the relationship between phonology and orthography, also known as feedback consistency) did not correlate with word recognition speed in a lexical decision task.

In addition to word frequency effects, word superiority effects have been cited as support for visual-route models. Word superiority refers to the phenomenon that actual words are judged to be words more quickly than non-words are judged to be non-words (Baluch & Besner, 1991; Tabossi & Laghi, 1992). According to a purely phonological-route model, as long as non-words contain permissible combinations of letters (e.g. 'plof' rather than 'nplf'), the reader should not know whether a word is a word or a non-word is a non-word until they reach the end of the word, or at least until they reach a point in the word where there are no more words that contain a particular combination of letters. If this were the process used to recognize words, non-words should be recognized faster than real words because a string of letters would be judged to be a non-word mid-way through the word, while identification as a real word could not be made until the entire string is processed. Thus, this rapid judgment of a string of text as a word or non-word requires addressed, whole-word visual-route processing rather than assembled, phonological-route processing.

Phonological Route

In strong phonological-route models, after the orthographic code is recognized, the related phonological code is activated. This activation then spreads to activate meaning nodes that match that phonological code. That is, the activation of phonology is a mediatory step in the recognition of words (see Figure 3). Two hypotheses have been proposed to explain how phonology is activated during word recognition. In one, the word is processed serially, activating the phonology of each letter and assembling the word from these pieces; this process is, therefore, called "assembled phonology". In the other type of proposed phonological activation, the word is viewed as a whole, much the same as the visual-route models propose, except that the phonology of the whole word is activated at once. Again, this occurs prior to lexical access; this hypothesis is called *addressed phonology*. Addressed phonology would seem to solve many of the discrepancies between visual- and phonological-route models (such as the frequency and word superiority effects described earlier). However, Lesch and Pollatsek (1998) showed that, while addressed phonology would solve many problems, this is not how people read. In this study, readers were shown a prime word (for example *pillow*), after which they judged whether the next word (for example 'bead') was semantically related to the first

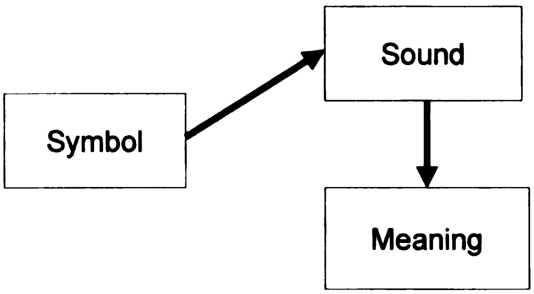


Figure 3. Phonology-mediated route to meaning (pre-lexical phonology).

word or not. Results showed that readers were sensitive to the similarity of the beginning of the word with the beginning of a word that is semantically related (to *pillow*, e.g. *be*ad and *be*d) as well as the possibility of the end of the word to be pronounced the same as a word that is semantically related (i.e. *-ead* as in *head* and the *ed* in *bed*). These results were interpreted as showing that the reader had processed the word piece-by-piece in an assembled manner, rather than addressed.

Strong phonological-route models are supported most strongly by research using a homophone interference paradigm, in which the reading of a printed word activates homophones of that word as well (Luo, 1996; Van Orden 1987). In Van Orden (1987), native English-speaking participants saw a category prime (such as *flower*) and had to judge whether a word (for example *rows*) fit the category or not. Similarly, in Luo (1996), participants were presented with a prime (such as *wolf*) and were asked which of two words (for example *lion* and *bare*) was semantically related to the prime word. In both of these studies, it was hypothesized that the phonology of the word presented would, due to the presentation of the prime, activate not only the word that was written (rows and bare), but also the exemplar word (e.g. *rose* and *bear*, respectively) and thus cause interference, resulting in slower and/or incorrect answers because of this confusion. Both of these studies did find longer reaction times and/or higher error rates with the homophone words compared to visually similar (and phonologically dissimilar) controls (for example *robe* and *bean*, respectively). Thus, these studies concluded that, in native English speakers reading in English, the phonology of words was activated prior to semantic activation (that is, phonological activation was pre-lexical).

Supporters of phonological-route models claim that the phonological route argument makes sense based on the fact that meaning is gleaned from spoken language using this same process, that of activating the phonological form and then the meaning (Perfetti, 2003). Considering the fact that in all languages around the world, the learning of spoken language precedes the learning of written language (Hoosain, 1991) and that, for nearly all people, encounters with aural/oral language is much more common than written language (Frost, 1998), it would stand to reason that the human mind is built to (or at least trained to) recognize meaning through recognition of sound. This is not to say that all written language has a direct one-to-one relationship with spoken language, but rather that because our brains already have an efficient system for using phonology to access the lexicon, it does not make sense to abandon that system in favor of a system that does not use phonology. At its most fundamental, written language is a derivative of spoken language, and most certainly not the other way around.

Although the findings that support phonological-route models do seem to be robust, the issues of word frequency and word superiority effects still remain. Aaron, Joshi, Ayotollah, Ellsberry, Henderson, and Lindsey (1999) addressed the existence of word frequency effects in a phonological-route model by showing that sight-word reading is strongly correlated with decoding ability; that is, the ability to recognize words as a whole may be dependent upon a prior ability to phonologically assemble those words. These findings are, however, dependent upon people whose decoding skills are weak to also be weak sight-readers. To the contrary, some studies have included participants who were strong sight-readers, but weak decoders (Freebody and Byrne, 1988). In addition, so far, no study has explained word superiority effects in terms of a phonological-route model.

It is due to the inability of either of these models to fully explain the word recognition process that dual-route models have emerged.

Dual Route

In response to the previously-discussed shortcomings of both strong visual- and strong phonological-route models, some researchers have proposed dual-route models that incorporate both a visual and a phonological pathway to the lexicon. There have been two separate interpretations of the term "dual-route" (cf. Van Orden, 1987). One identifies two separate routes that occur not in parallel, but rather in a serial, cascading manner, with the result of visual-route processing feeding into phonological-route processing. In essence, these are only visual-route models with a phonological route tacked onto the end. On the other hand, most dual-route models specify two separate routes, occurring in parallel, and it is this type of dual-route model that will be discussed here.

Dual-route models have the advantage, due to the incorporation of both a phonological and visual route running in parallel (Figure 4), of being able to account for both homophone and word frequency effects. Dual-route models are frequently presented as a metaphorical horse race (e.g. Paap & Noel, 1991) in which a phonological horse and a visual horse are racing toward a common goal, lexical activation. In most of these models, the horses do not interact, but rather the fastest horse (route) to reach the lexicon wins. Dual-route models where the horses do interact have also been proposed (e.g. Seidenberg, 1985; Seidenberg, 1992). These models think of activation in terms of

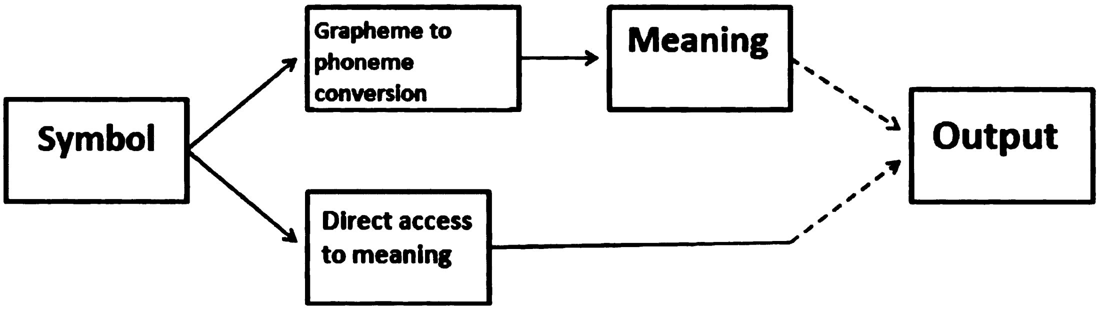


Figure 4. Dual-route model.

a threshold; information is gained from a combination of phonology and orthography until a threshold is reached, resulting in lexical activation.

Orthographic Depth Hypothesis

Somewhat of a product of the dual-route models is the orthographic depth hypothesis (ODH) (Frost, 1994; Frost & Katz, 1992; Frost, Katz, & Bentin, 1987). While the ODH does incorporate both reliance on visual information and phonological information, it discusses phonological activation not as a binary question of one route or another, but rather in terms of degree of activation. The ODH proposes that the extent to which the written form of a language corresponds with the phonological form determines the process used to recognize words (Akamatsu, 2005; cf. Frost, 1994). That is, the more orthographically shallow a language is, the greater extent phonology will be activated during word recognition; on the other hand, more orthographically deep languages will involve less phonological activation. In a study by Katz and Feldman (1981), it was found that word naming depended more on phonology in Serbo-Croatian (an orthographically very shallow language) than in English (orthographically deeper than Serbo-Croatian), and word naming depended more on phonology in English than in Hebrew (orthographically deeper than English). Similar effects have been demonstrated in a number of studies in a variety of languages (Welsh: Ellis & Hooper, 2001; Hebrew: Geva & Siegel, 2000; English, French, and Spanish: Goswami, Gombert, & De Barrera, 1998; Turkish: Öney & Durgunoglu, 1997; 13 European languages: Seymour, Aro, and Erskine, 2003)

Regarding the languages involved in this study, the ODH indicates that, for native speakers of the languages, reading in English would involve a much greater degree of phonological activation than reading in Chinese. This is due to the existence of GPC rules in English, and the lack of a reliable phonological cue in Chinese characters. This is not to say that phonological activation does not occur while reading Chinese characters; on the contrary, there have been a number of studies by Perfetti and colleagues that demonstrate phonological activation during reading in Chinese (e.g., Perfetti & Tan, 1998; Perfetti & Zhang, 1995; Tan & Perfetti, 1997, 1998, 1999; Tan, Spinks, Gao, Liu, Perfetti, Xiong, Stofer, Pu, Liu & Fox, 2000). While these studies do claim that phonology is activated during word recognition in both English and Chinese, they also state that whereas word recognition in English can involve cascaded activation of phonology (which suggests assembled phonology), in Chinese, "phonology awaits the completion of orthographic processes" (Perfetti et al, 2007, p. 133), suggesting addressed recognition. The general conclusion of these studies is that, in any language, phonology is activated to the extent "allowed by the writing system" (Perfetti, Zhang, & Berent, 1992, p.245). This statement would tend to be in agreement with the ODH in that the ODH also states that the degree of phonological activation depends on the writing system of the language.

Transfer

The degree to which phonological information is used affects more than just the first language. Much research has shown that L2 learners' word recognition is heavily influenced by the way that words are recognized in their L1 (Brown & Haynes, 1985; Chikamatsu, 1996, 2006; Green & Meara, 1987; Koda 1988, 1989, 1990; Mori, 1998; Wade-Woolley, 1999; Wang, Koda, & Perfetti, 2003). Such transfer includes the degree to which phonological information is activated while reading. Koda (1989) investigated L1 Japanese (a relatively deep orthography that is based on Chinese) learners of English (which is comparatively much more shallow) and showed that these learners could recall pronounceable strings of letters no better than unpronounceable strings of letters, and in fact could sometimes recall the unpronounceable strings better. That is, they did not make use of the phonological information that was available in order to make the letters into words (albeit nonsensical words), and instead remembered the visual representation of the letters. These results were interpreted as being a result of transfer from the L1 recognition process, which, according to the ODH, includes a lesser involvement of phonology.

Mori (1998) investigated L1 English and L1 Chinese learners of Japanese in a similar fashion. In this study, participants were presented with pseudoword Japanese kanji characters, some of which were embedded with hiragana (Japanese syllabary) characters. In a recognition task, L1 Chinese participants were able to recall an equal number of pseudocharacters that had embedded phonologically-accessible elements (the hiragana characters) as those that did not contain such elements. On the other hand, the L1 English learners remembered a significantly greater number of pseudoword kanji that had embedded pronounceable parts than they did kanji that had no such element, thus showing that the L1 English learners made use the sound element of the characters, whereas the L1 Chinese did not. This was, again, interpreted to be an effect of the L1 word recognition process being transferred to the L2. Such a strategy would be more similar to the way L1 English speakers would process English.

Although these studies showed L1 transfer of the use of phonology, they showed it via offline recall tasks, which may be different from online reading. Chikamatsu (1996) also investigated L1 English and L1 Chinese learners of Japanese, and found similar results to the previous studies, although through a different methodology. This study involved a script-manipulation task using the syllabic characters of Japanese (hiragana and katakana). In this task, a number of words that students were more familiar with seeing in hiragana were written in katakana and vice versa, thereby preventing visual recognition of the words and forcing use of phonology. She found significant effects of visual familiarity (that is, the script that the words were written in) and also word length, with familiarity affecting the L1 Chinese group more than the L1 English group, thus showing that the L1 English learners were not as disturbed by changes in the visual appearance of the words as long as phonological information did not change. In addition, word length affected the L1 English group more than the L1 Chinese group, showing that, for the L1 Chinese group, word length did not make as much of a difference, which would be the result of addressed visual recognition.

In a study involving L1 Chinese and L1 Korean learners of English, Wang, Koda, and Perfetti (2003) investigated learners' relative reliance on phonological and visual information during word recognition. Chinese was chosen because Chinese orthography, according to the ODH, should involve a minimal amount of phonological activation, while Korean was chosen because although Korean and Chinese are both syllabaries, Korean orthography (*hangul*) is alphabetic in nature and thus, according to the ODH, should involve more activation of phonology (Cho & Chen, 1999). This study involved a semantic category judgment task, similar to that developed by Van Orden (1987), in which, after the presentation of a category prime (such as *a flower*), there were foils that were phonologically-similar (e.g. *rows*) or controls that were orthographically-similar (e.g. *robe*) to actual category exemplar words (e.g. *rose*), thus allowing for a measure of relative use of phonological information and orthographical information during L2 word recognition. In addition, to measure reliance on orthographical information, the orthographic similarity to exemplar words was also manipulated. In this task, L1 Korean learners of English were found to make more false-positive errors with regard to the phonologically-similar foils than the orthographically-similar foils, whereas the L1 Chinese learners did not make more errors on one than the other. While this does suggest that there is phonological activation for the L1 Chinese learners, there is less than for the L1 Korean learners. This difference was theorized to have been a result of transfer from the learners' L1 processing. To confirm these findings, a phoneme deletion task was also conducted. In this task, L1 Chinese learners performed more poorly than the L1 Korean learners, and also made more errors which, although they were orthographically acceptable, were phonologically incorrect.

Proficiency

Many of the studies described above were conducted with intermediate to advanced L2 learners. Very few studies have been conducted with lower-level L2 learners, and even fewer have been conducted comparing low proficiency and high proficiency learners. A number of studies did not operationalize proficiency at all. Wang, Koda, and Perfetti (2003) state that the results of previous studies investigating L1 transfer of word recognition skills are "clouded by the lack of attention to the level of English proficiency in the L2 groups," and further state that it is "important to separate whether differences arise from the L1 writing system background or from the level of skill attained in L2" (p. 134). In their study, the majority of participants were recruited from intermediate ESL courses, with additional participants recruited from advanced ESL classes and among first-year graduate students. In their description of their results, which support the ODH and L1 transfer, Wang et al. further state, "Moreover, the logographic L1 transfer effect on alphabetic L2 learning may occur only at the beginning stages of learning. With increasing proficiency in English, the effect of phonology on English word processing will eventually prevail in Chinese L2 readers' performance" (p.144). As strong of a statement as this is, proficiency was not a variable in this study; rather it was controlled due to the possibility of effects of proficiency on the degree to which phonology is used during word recognition.

One important element to consider in research investigating changes in word recognition is whether changes occur due to actual restructuring of the learner's cognitive processes, or if changes are simply due to higher proficiency learners processing words in the same way, but faster than lower proficiency learners. Due to increased proficiency, it is possible that better experimental results may be due to a speeding up of performance, rather than actual restructuring of cognitive processes, leading to automatization. It is important to include analyses of automatization in word recognition studies because when learners' word recognition processes are automatized, fewer cognitive and attentional resources are consumed (Segalowitz & Hulstijn, 2005).

This phenomenon was investigated by Segalowitz and Segalowitz (1993) and Segalowitz, Segalowitz, and Wood (1998). They implemented a statistical procedure using the learners' reaction times (RT) and standard deviations of those reaction times (SDrt) to determine a coefficient of variation (CVrt). In the case of simple speed-up, mean RT and SDrt decrease proportionally, which results in no change in CVrt. On the other hand, when improvement in performance is due to increased efficiency of processing (i.e. automatization), SDrt will decrease to a greater extent than mean RT, thus resulting in a reduced CVrt (Akamatsu, 2007). Incorporating this procedure, Segalowitz and Segalowitz (1993) using L1 French and L2 English, and Segalowitz, Segalowitz, and Wood (1998) using L1 English and L2 French determined that learners' L2 processing was, in fact, restructured as proficiency increased. Incorporating this concept into cross-orthographic word recognition research involving intra-word processing, Akamatsu (2007) found that, with practice, the word recognition processes of L1 Japanese learners of English were restructured, although only in the case of low-frequency words; in the case of high-frequency words, simple speed-up was shown.

Chikamatsu (2006) also took the difference between speed-up and automatization into account during her investigation of the use of phonology during word recognition of L1 English learners of Japanese. This study used a similar script manipulation paradigm as Chikamatsu (1996), described earlier. A lexical decision task using decontextualized words was used as well as a task where participants read paragraphs containing words in changed scripts. The changed script forced the participants to recognize words not by their overall visual appearance (a more visual route), but rather to process the words in a serial manner, assembling the phonology of the word. In these experiments, it was found that advanced students (fourth semester university students of Japanese) were more affected by visual unfamiliarity than beginning students (second semester) in the decontextualized lexical decision task. In addition, this was shown to be a restructuring of the word recognition process, not merely speed-up, although no statistics relating to this were reported. These results were interpreted as showing that as learners become more proficient in a language, the cognitive processes of word recognition adapt to the L2.

While the previous study seems to suggest that with increased proficiency, word recognition processes are indeed restructured, conflicting evidence has also been shown. Akamatsu (2005) also investigated the effect of proficiency on the use of phonology during the word recognition processes, although using L1 Japanese learners of English. Similar to Chikamatsu (2006), this study used a type of script manipulation. In this study, alternating case in English (e.g. aLtErNaTiNg CaSe) was used in a word naming task. Akamatsu (2005) argues that this paradigm works well as a measure of L1 orthographic effects because readers from L1s with non-phonological orthographies (e.g. Chinese and Japanese) show a large impact in ability to process the constituent letters of words (Brown & Haynes, 1985). Following the ODH, the processing of the individual letters of words would indicate activation of the phonology of those letters (i.e., assembled phonology). In this experiment, expected results were operationally the opposite of Chikamatsu (2006), due to the languages involved, although paradigmatically the same. That is, although case manipulation was expected to have a greater effect on the lower proficiency group, which is opposite of Chikamatsu (2006), such a result would indicate a change in processing to adapt to the script used in the L2, which is the same as Chikamatsu hypothesized. However, the results showed that the high-proficiency English learners did not perform any better than the low-proficiency group. This was interpreted to mean that the effects of the L1 orthography on L2 word recognition were so profound that any amount of increase in proficiency and/or exposure (the high-proficiency group was also in an ESL context while the low-proficiency group was in an EFL context) could not reduce them (Akamatsu, 2005).

The conflicting results of these studies show that more research is necessary in order to determine the precise effect of proficiency on cross-orthographic word recognition.

Research Questions

The various studies described have informed the present study, and have led to the following research questions:

1) Does L2 proficiency affect L2 readers' processing of the constituent letters of words?

   a) If so, do differences in readers' abilities to process constituent letters of words depend on the frequencies of the words?

   b) If so, do differences in readers' abilities to process constituent letters of words depend on the spelling regularity (i.e., regularity of GPC rules) of the words?

2) When high L2 proficiency readers read English words for meaning, do they show more reliance on phonological information compared to low proficiency readers? Do lower L2 proficiency readers show heavier reliance on orthographic information than high L2 proficiency readers?

Hypotheses

Regarding the first research question, the processing of the constituent letters of words was chosen as a measure of phonological involvement as previous studies (Akamatsu, 1999, 2005) used this methodology to measure the same aspect of reading. Intra-word processing is an area in which L2 readers whose L1s have non-phonological orthographies often have a greater amount of L1 transfer (Brown & Haynes, 1985; Holm & Dodd, 1996; Koda, 1999; Wang, Koda, & Perfetti, 2003); that is, because their L1 orthography does not allow for the concept of intra-word processing and assembled phonology, L2 English readers whose L1 orthography is non-alphabetic often have trouble with this part of reading English words. In the present study, it is hypothesized that L2 proficiency will have an effect on L2 readers' processing of the constituent letters of words. Based on the research described previously, a high level of L2 reading proficiency should reduce the effect that the L1 orthography has on L2 word recognition. It is hypothesized that the effect of the L1 orthography will be greater for lower-proficiency L2 readers than for higher-proficiency L2 readers.

Previous research has shown that word frequency does have an effect on word recognition processes in both L1 and L2 (i.e., more frequent words are more often processed via an addressed visual route, whereas less frequent words are more likely to be processed via an assembled phonological route) (Forster & Chambers, 1973; Hino & Lupker, 2000); in addition, previous research has shown that words that do not follow regular grapheme-phoneme correspondence rules (termed *exception spelling words* in the present study) are more likely to be processed via a visual route whereas words that do have regular GPC rules are more likely to be processed via a phonological route (Hino &

Lupker, 1998; Plaut, McClelland, Seidenberg, & Patterson,1996). For those reasons, it is hypothesized that word frequency and spelling regularity may have an effect on the degree to which L1 orthography affects L2 word recognition.

In response to the second research question, it is hypothesized that high proficiency readers will show more reliance on phonological information than low proficiency readers. This is because it is hypothesized that as proficiency increases, learners' word recognition processes will be restructured to adapt to the L2 script. In the context of the present study, according to the ODH, the script of English involves a greater degree of phonological involvement than Chinese, and as such, high L2 English proficiency readers whose L1 is Chinese will show greater reliance on phonology than low L2 English proficiency readers, and low L2 English proficiency readers will rely more on orthographic information than phonological.

Methodology

General Procedure

This study sought to investigate the effect of L2 proficiency on L2 word recognition processes. This was achieved by measuring readers' ability to process the constituent letters of words and readers' reliance on the phonology of words. According to the orthographic depth hypothesis, reading in Chinese involves a lesser degree of phonological activation, whereas reading in English involves greater activation; this study was concerned with potential differences in the amount of phonological activation at different levels of proficiency. In order to achieve this, participants' proficiency was measured at the time of testing via administration of portions of the Michigan Test of English Language Proficiency (MTELP). In addition, participants completed a background questionnaire. Two computerized experiments were conducted with all participants; the order in which participants performed the experiments and the order of the items within all experimental tasks were counterbalanced.

During recruitment, all participants were given an information sheet (Appendix A) which gave a general description of the tasks, benefits, and risks associated with participating in the study, and contact information for the investigators. Because the study involved some participants whose proficiency in English was expected to be low, an attempt was made to make the form as simple and as clear as possible, while still including necessary information. In addition, the consent form was exactly the same as the advertising flyer that participants received prior to participation, and as such, participants had a chance to read the form at their own pace before coming to participate in the study. Participants agreed orally to the consent form[2]. All participants received compensation of $15 cash for participating in the study, after which4 they signed a statement saying they had received the compensation. This form was, however, stored separately from each participant's data in order to protect participants' privacy. Participants were identified only by a subject ID number on all test materials and analyses, except the signed statement of receipt of compensation; the subject ID numbers were used only to correlate data from different data sets.

Each participant attended a single 90-minute session arranged with the author. During this time, all participants completed: 1) consent form, 2) background questionnaire (Appendix B), 3) proficiency test, 4) Experiment 1(lexical decision task)

[2] A signed consent form was included in the original Institutional Review Board application. However, the IRB reviewer recommended that the advertisement flyer be used as the consent form instead of the original signed form, as the advertisement flyer contained all of the information appropriate for an 'Exempt' status study, and by not having participants sign a consent form, the form would not need to be kept for the required three years.

and Experiment 2 (semantic category judgment task) (the order of the experiments was counterbalanced), and 5) debriefing form (Appendix C) and compensation for participation. The following will describe the background questionnaire used as well as the proficiency measure, after which the two experiment will be described in detail.

Background Questionnaire

After agreeing to the consent form, all participants completed a two-page background questionnaire (Appendix B). Approximately 12 minutes were allotted for completion of the background questionnaire. The questionnaire was partially based on the pretest version of the Language Contact Profile by Freed, Dewey, Segalowitz, and Halter (2004). The questionnaire included questions about L1, gender, age, and current and previous L1 and L2 use and exposure.

In addition, the background questionnaire asked about the participant's use of *pinyin*, the method of learning the pronunciation of Chinese characters through the use of the roman alphabet. This system is used mainly in mainland China, and is taught to native Chinese-speaking children from a young age (usually when they enter first grade in elementary school) and is used when teaching Chinese characters in school. This system was adopted by the government of China in 1958, and has been used in schools since then (Yin & Felley, 1990). In Taiwan, this system is not used, and instead, the sounds of Chinese characters are shown by using a specific set of other Chinese characters. A number of studies have shown that the increased exposure to the roman alphabet through pinyin can have an effect on the way Chinese speakers recognize English words (e.g. Everson, 1998; Tan & Perfetti, 1998; Wang & Geva, 2003). Some participants from Taiwan did ask about this question, as the system used in Taiwan is also sometimes called *pinyin*, although the majority of Taiwanese participants did understand pinyin to mean the romanized variant. The system that is most widely used in China is called *hanyu pinyin*, and it may be possible that this confusion could have been eliminated by using this term instead of only *pinyin*. This question was purposely left as only *pinyin* in order to include other systems of romanization that are occasionally used throughout China (such as *tongyong pinyin*).

Proficiency Measure

After completing the background questionnaire, participants took the grammar and reading portions of the Michigan Test of English Language Proficiency (MTELP), form P, acquired from the University of Michigan English Language Institute. This test was chosen because it has a long history of use in language acquisition research as well as as a valid and reliable placement test in language programs (Baldauf & Dawson, 1980). In addition, it was chosen because participants were unlikely to have previous exposure to the test, as it is not used by Michigan State University to test language proficiency, nor is it accepted by the university as proof of language proficiency. It was decided to administer only these two sections of the test because of time constraints, and the vocabulary section provided only a minimal amount of information regarding the participants' proficiency. As participants took two out of three sections of the test, the total amount of time allowed to take the test was reduced proportionally. The University of Michigan English Language Institute recommends that a total of 75 minutes be allowed to complete the entire test; therefore, participants in the current study were allowed 50 minutes, which is two-thirds of the full time to complete two-thirds of the full test.

Participants

All participants were native speakers of Chinese enrolled at Michigan State University in one of the following: the MA program in Teaching English to Speakers of Other Languages, the Second Language Studies Ph.D. program, Level 2 or Level 3 classes at the English Language Center (ELC), or Level 2 of a community outreach English as a Second Language course taught at Michigan State University. Participants were contacted either by email (MA TESOL and SLS Ph.D. students) or in their classes (ELC and community outreach course students), with institution and instructor approval. While age restrictions were not reported in many of the studies closely related with this study (Akamatsu, 2005; Chikamatsu, 2006; Wang, Koda, & Perfetti, 2003), the participants included in those studies were undergraduate university students, who would most likely be from a much more limited age group. Because the present study included a number of graduate students as well as community members taking the community outreach course, an age restriction was instituted. Following a number of psycholinguistic studies (e.g. McCann, Folk, & Johnston, 1992; Rumsey, Horwitz, Donohue, Nace, Maisog, & Andreason, 1997), all participants included in this study were between 18 and 40 years of age. All participants were from either mainland China or Taiwan. Participants' status as native speakers of Chinese was confirmed by the background questionnaire. All participants received $15 cash for their participation.

In total, 34 participants completed all parts of the study. Although participants were recruited from populations that were expected to yield distinct high and low proficiency groups, upon inspection of the results of the proficiency measure, it was discovered that this was, in fact, not what was shown; rather, proficiency scores were fairly continuous between the lowest score of 19 and the highest score of 55 (of a maximum of 60) (Figure 5). Because this study sought to investigate the effect of proficiency differences on word recognition, it was imperative that proficiency differences be evident between the two groups. For this reason, it was decided to implement an extreme groups approach (EGA). The 34 participants were divided into thirds (a high proficiency group, a medium proficiency group, and a low proficiency group), which yielded 11.3 participants per group. Because portions of this study implemented two word lists, for the purpose of counterbalancing it was necessary that there be equal numbers of participants who saw each list in each group, thus necessitating an even number of participants in each group. Therefore, groups were limited to 10 participants each. The 10 participants with the highest total MTELP scores were labeled

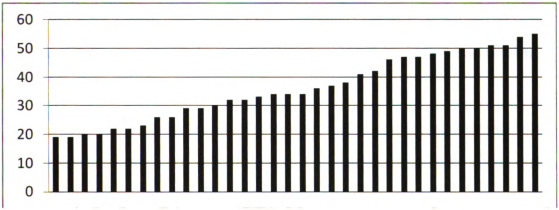


Figure 5. MTELP scores for all 34 participants (maximum possible score of 60).

as the high proficiency group; scores in this group ranged from 47 to 55 out of 60. The 10 lowest scores were labeled as the low proficiency group; scores in this group ranged from 19 to 29. Only the outer tertiles were included in the analysis. Although the use of EGA is controversial because it produces artificially inflated effect size estimates, it is beneficial when the exact relationship between the variables involved is unknown, and the goal of the research is to determine whether an interaction exists, regardless of the magnitude (Preacher, Rucker, MacCallum, & Nicewander, 2005); it is for this reason that it was felt that EGA was justified in the context of the current study. However, because a result of the use of EGA is artificially inflated effect size estimates, these will not be reported in the present study.

In addition to MTELP scores, as a part of the background questionnaire, all participants rated their own abilities (in both Chinese and English) in the categories of listening, speaking, reading, writing, and grammar (Table 1). Self-evaluations have also been collected in other psycholinguistic studies (e.g. Jiang, 2007; Midgley, Holcomb, & Grainger, 2008), and can often provide insight into or corroboration of tested measures. While all participants rated themselves similarly as native speakers of Chinese, the self-ratings of the high-proficiency group were significantly higher in all categories (listening, speaking, reading, writing, and grammar) than the self-ratings of the low-proficiency group.

The number of male and female participants was the same between proficiency groups (female n = 8, male n = 2). There were more pinyin users in the low-proficiency group (n = 8) than the high proficiency group (n = 6). More participants in the high proficiency group reported using some English at home (n = 4) than in the low proficiency group ($n = 1$), and the same was seen in terms of pre-university education in English (high proficiency $n = 3$; low proficiency $n = 1$). The mean age was higher for the high proficiency group than for the low proficiency group although this difference was not significant (Table 2). Length of residence in English speaking places (including Michigan and any other primarily English-speaking locations) was significantly greater for the high proficiency group than for the low proficiency group, although it seems logical that those who have been immersed in the L2 environment would have higher proficiency levels.

Table 1

Means (and standard deviations) of proficiency measure results

Prof.	Self-reported proficiency scores					MTELP Scores		
Group	Listening**	Speaking*	Reading***	Writing**	Grammar*	Grammar***	Reading***	Total***
High	3.65 (.747)	3.15 (.818)	3.95 (.599)	3.45 (.762)	3.60 (.843)	34.20 (2.39)	16.00 (1.16)	50.20 (2.70)
Low	2.40 (.699)	2.30 (.675)	2.40 (.843)	2.20 (.633)	2.80 (.633)	15.80 (2.82)	6.80 (2.20)	22.60 (3.41)

* Group means were significantly different at $p<.05$ level

** Group means were significantly different at $p<.01$ level.

*** Group means were significantly different at $p<.001$ level.

Table 2

Means (and standard deviations) of participant age and length of residence in English-speaking locations

	Age in years	Residency in months*
High Proficiency	28.40 (4.90)	25.83 (19.37)
Low Proficiency	23.10 (7.19)	8.60 (4.97)

* Group means were significantly different at $p < .05$ level

Experiment 1

Experiment 1 implemented a lexical decision task in order to address research question 1: "Does L2 proficiency affect L2 readers' processing of the constituent letters of words?", as well as the sub-questions: "If so, do differences in readers' abilities to process constituent letters of words depend on the frequencies of the words?" and "If so, do differences in readers' abilities to process constituent letters of words depend on the spelling regularity of the words?" In a lexical decision task, a participant is presented with a single word and must respond as quickly and accurately as possible whether that word is a word in English or not. Research question 1 was addressed by manipulating the case in which the stimuli words were written, which forced readers to process constituent letters in the stimulus words manipulated in this way. The two sub-questions were addressed by manipulating the frequency and spelling regularity of the stimuli words (also described below).

Materials

Participants completed a lexical decision task that contained a total of 160 items, including 80 critical words and 80 distractor words (Appendix D); in this task, participants were to choose whether the word presented was a real English word or not. Case alternation (e.g. cAsE aLtErNaTiOn) was implemented in this experiment in order to force readers to process words in an assembled, serial way while reading the manipulated words. Half of the words were presented in this visually unfamiliar form, and half of the words were presented in a visually familiar form (all lower case). This methodology has been used in a number of studies investigating L2 word recognition (Akamatsu, 1999, 2003, 2005; Kinoshita, 1987) and L1 word recognition (Ehri & Wilce, 1982), as well as other word-shape effects on reading (e.g. Fisher, 1975). As in Akamatsu's studies, by presenting the words in a visually unfamiliar form, visual cues become inaccessible to the reader, making a visual route of word recognition impossible. The serial parsing of a word is theorized to activate the phonology of each letter (Paap, Newsome, & Noel, 1984). Word case was included as a within-subject independent variable.

Two word lists were used in this experiment; each list contained the same words, but words that were in lower case in one list were in alternating case in the other list, and vice versa. Half of the participants in each group saw List 1 and half of the participants saw List 2. This was to avoid any effect of the words used; that is, if only one list were used, it is possible that the list of lower case words and the list of alternating case words may have had inherent differences. By using two word lists containing the same words but in different cases, the possibility of inherent differences is removed because the lower case words for one group and the alternating case words in the other group have exactly the same attributes, except for the case in which they are presented.

In addition, previous studies have shown that word frequency and spelling regularity have significant effects on word recognition in both L1 English speakers (e.g. Hino & Lupker, 2000; Jared, 2002) as well as L2 English speakers (Akamatsu, 1999), and as such, those factors were also included as word-based variables.

Differences between the critical words were operationalized in terms of frequency, spelling regularity, and letter case; length and word class were controlled (Table 3). Frequency was determined using Kucera and Francis (1967). Because the Kucera and Francis (1967) corpus was compiled from native-speaker data, and it was assumed that English learners most likely had had different exposure to language than native speakers, stimuli were selected based first on frequency, and secondly on their appropriateness to L2 speakers of English based on the author's personal experience as an ESL/EFL teacher. The stimulus words that were used in the experiment, however, still follow the frequency counts determined using Kucera and Francis (1967). Frequency between the high frequency conditions and the low frequency conditions was significantly different as shown by *t*-test, $t(39) = 11.99, p<.0001$.

Spelling regularity was determined using Berndt, Reggia, and Mitchum's (1987) study of grapheme-phoneme correspondence probability as a guide. The regular spelling words had grapheme-phoneme probabilities that were above 70% (or the highest probability for that grapheme), whereas the exception spelling words contained grapheme-phoneme probabilities of 30% or less, although most were below 10%. Although a number of studies have used Venezky (1970) to determine spelling regularity (e.g. Hino & Lupker, 2000; Holm & Dodd, 1996), Venezky's rules are not based on quantitative evidence (Kessler & Treiman, 2001), and as such, Berndt et al. (1987) was used instead (and therefore the term "exception spelling" is used rather than "irregular spelling"). In addition, Akamatsu (1999, 2005), on which this study is based, also used Berndt et al. (1987) to make spelling regularity classifications.

Words in each condition were matched for length, with word length differences being non-significant between each of the word factor conditions. Word class was matched within frequency levels (and therefore between spelling regularity levels), however there were differences in word class numbers between frequency levels; the rarity of low frequency adjectives and adverbs that are between three and five letters in length and that that show extreme spelling regularity patterns contained in the Kucera and Francis corpus necessitated this difference.

In addition, a number of the stimuli from Akamatsu (1999, 2005) were, with permission, included. The stimuli set for those studies had also been categorized using Berndt et al. (1987) for spelling regularity and Kucera and Francis (1967), among other corpora, for frequency. However, because the current study necessitated twice the number of stimulus words as Akamatsu's, the stimuli from that study were mixed in among other stimuli determined by the method described above. The stimuli from Akamatsu's studies were included with their frequencies as shown in Kucera and Francis (1967).

Non-word distractors were constructed by changing the first 1-2 letters of common English words not used elsewhere in the stimuli. Because of this, the spelling patterns of the non-word distractors followed common English spelling patterns. Half of the non-word distractors were presented in alternating case and half were presented in lower case.

Table 3

Descriptive statistics for frequency, length, and word class for each condition

	Frequency (OPM)		Length (letters)					
Condition	*M*	*SD*	*M*	*SD*	Nouns	Verbs	Adj.	Adv.
High-Reg	401.70	249.74	4.15	0.49	8	6	4	2
High-Exc	382.05	162.69	4.25	0.44	8	6	4	2
Low-Reg	10.45	6.98	4.20	0.62	12	6	2	0
Low-Exc	10.35	6.46	4.25	0.72	12	6	2	0

Note: OPM = Occurrences per million, as determined using the Kucera and Francis (1967) corpus; High and Low refer to frequency categorization, and Reg and Exc refer to spelling regularity categorization as either regular or exception spelling, according to Berndt et al. (1987)

Procedure

The task was conducted using the DMDX software package (Forster & Forster, 2003) on Windows XP-based computers. Participants were seated approximately 50cm from the monitors, with blinder walls on either side of the participant. Stimuli were presented in 24-point font in black text on a white background. After reading the instructions (Appendix E), each participant completed ten practice items, of which five were presented in lower case and five were presented in alternating case, and five were real words and five were non-word distractors (all practice items were, however, presented in a randomized order). Before each stimulus item, participants were presented with a fixation mask, which consisted of five hash marks (#####), for 500ms. Participants then made a decision as to whether they thought the presented word was a real word in English or not. Participants responded using their left and right index fingers.

To respond that a word was a real English word, participants pressed the 'L' key on the keyboard (covered with green paper to indicate a positive response), and to respond that a word was not a real word in English, participants pressed the 'D' key on the keyboard (covered with red paper to indicate a negative response). Both response time and accuracy were recorded for each word. Participants were given a break after half (80) of the items. Participants were asked about their level of fatigue after each task and informed that they could take a break if they needed to; however, no participants responded that they were fatigued.

Results

The analysis utilized a 2 x 2 x 2 x 2 (two levels each of participant L2 proficiency, stimulus word frequency, stimulus word spelling regularity, and stimulus word case) repeated measures factorial design. Data were analyzed both by participant and by item.

In the participant analysis, L2 proficiency was the between-subjects independent variable and the word conditions (frequency, regularity, and case) were within-subject independent variables. In the item analysis, L2 proficiency and case were within-item factors, and spelling regularity and frequency were between-item factors. Dependent variables in both analyses were reaction time and error rate.

In order to preserve data, responses greater than two standard deviations from the mean of each participant were replaced with a value equal to two standard deviations from the mean rather than discarded. Trimming data at two standard deviations from the mean of each participant in each condition is a common procedure in word recognition research (e.g. Holmes, 1998; Jiang, 1999). The trimming of data affected 2.34% of the data. The analyses of the reaction time data were based only on correct responses. Means for reaction times and error rates are presented in Table 4.

Reaction Times

Overall, there was a main effect of proficiency group, both by participant, $F(1,18) = 11.30, p < .01$, and by item, $F(1, 36) = 176.03, p < .001$, showing that, on the whole, the more proficient readers recognized words faster than the less proficient readers.

In addition, there was a main effect of case, with participants recognizing lower case words significantly faster than alternating-case words, by participant, $F(1,18) = 59.72, p < .001$, and by item, $F(1,36) = 77.74, p < .001$. There was also a main effect of word frequency; high frequency words were recognized more quickly than low frequency words, both by participant, $F(1,18) = 99.40, p < .001$, and by item, $F(1,36) = 52.13, p < .001$.

The main effect of spelling regularity was not significant by participant or by item. However, there was a significant three-way interaction between spelling regularity, case, and proficiency that was shown only in the item analysis (Figure 6), $F(36) = 4.98, p < .05$. Post-hoc paired-sample t-tests showed that the effect of spelling regularity between the two proficiency groups was significant only in the lower case condition, $t(19) = 2.46, p < .05$.

There was a significant interaction between frequency and proficiency group by participant, $F(1,18) = 8.59, p < .01$ and by item, $F(1,36) = 26.68, p < .001$, showing that word frequency had more of an impact on the low proficiency group than it did on the high proficiency group.

Table 4

Means (and standard deviations) of reaction times and error rates

	Reaction Times (ms)		Error Rates (%)	
Condition	High Prof. *M* (*SD*)	Low Prof. *M* (*SD*)	High Prof. *M* (*SD*)	Low Prof. *M* (*SD*)
Alternating case (visually unfamiliar)				
High, Reg.	888.62 (142.7)	1064.00 (157.0)	5.00 (6.3)	3.12 (4.8)
High, Exc.	854.19 (132.3)	996.88 (146.2)	1.88 (5.4)	4.64 (5.5)
Low, Reg.	1176.75 (214.4)	1392.75 (336.0)	10.07 10.6)	32.63 (21.2)
Low, Exc.	1072.12 (190.0)	1371.25 (240.3)	5.07 (6.4)	35.07 (17.5)
Lower case (visually familiar)				
High, Reg.	727.88 (110.2)	775.44 (116.9)	0.62 (2.5)	0.00 (0.0)
High, Exc.	686.88 (77.5)	818.75 (121.6)	0.00 (0.0)	1.88 (5.4)
Low, Reg.	903.81 (108.1)	1023.06 (130.3)	10.38 (10.7)	21.25 (16.3)
Low, Exc.	859.12 (151.1)	1068.69 (131.3)	9.38 (16.5)	28.53 (12.6)

Note: 'High' and 'Low' refer to frequency as determined using Kucera and Francis (1967); 'Reg.' and 'Exc.' refer to regular spelling and exception spelling, as determined using Berndt et al. (1987).

Most directly related to research question one, however, was the interaction between proficiency and case. The results did reveal a significant interaction between case and proficiency group, by participant: $F(1,18) = 5.47, p < .05$, and by item: $F(1,36) = 11.85, p < .005$ (Figure 7). However, this may be qualified by a significant three-way interaction between proficiency group, spelling regularity, and case, significant in the item analysis, $F(1,36) = 4.98, p < .05$, but non-significant in the participant analysis. Post-hoc paired-sample *t*-tests showed that the effect of alternating case text between the two proficiency levels was significant in the regular spelling word condition, $t(19) = -4.01, p < .001$, but non-significant in the exception word condition, $t(19) = -.875, p > .3$. This difference can be seen in Figure 8.

Even though the interaction between proficiency and case was significant, the question remains about whether this effect is due to speed-up (i.e., more proficient readers just read faster), or whether actual restructuring is occurring. As discussed earlier, Segalowitz and Segalowitz (1993) developed a statistical measure of automatization to investigate differences between speed-up and restructuring. This method has been subsequently been used in a number of studies investigating restructuring of cognitive processes (e.g. Akamatsu, 2007; Chikamatsu, 2006; Fukkink, Hulstijn, & Simis, 2005; Jiang, 2007). According to Fukkink et al. (2005), "In the case of true automatization… component processes become more routinized or are eliminated altogether. Hence, not only will mean RT and SDrt reduce, but so will CVrt" (p.57). CVrt is the coefficient of variation of the reaction times, which is found by dividing the standard deviation of the reaction time by the mean reaction time. The data from the present study are presented in Table 5. The mean reaction time, SDrt, and CVrt all decrease from low proficiency to high proficiency readers. Paired-sample *t*-tests on the CVrt of low proficiency and the CVrt of high proficiency participants showed significant differences both in the case of exception recognition process, and not merely speed-up.

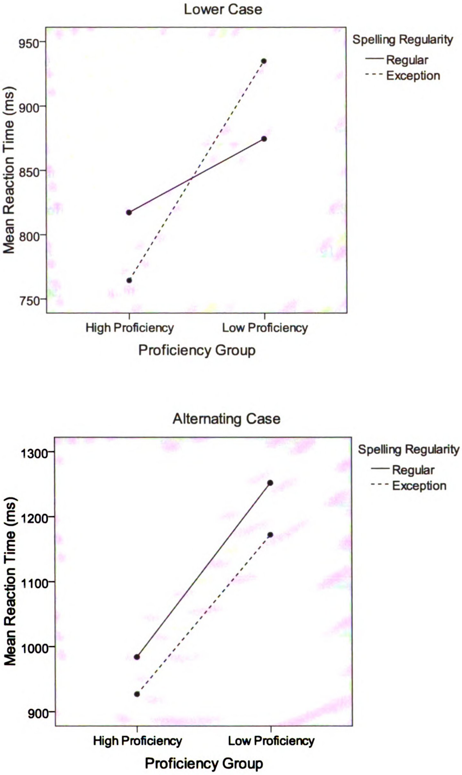


Figure 6. Interaction between spelling regularity and proficiency within case.

Table 5

Summary of means, standard deviations, and coefficients of variation for alternating case stimuli

	Mean RT	SDrt	CVrt
Both Spelling Conditions			
Low Proficiency	1343.26	688.81	.4941
High Proficiency	951.48	291.85	.3051
Regular Spelling Conditions			
Low Proficiency	1369.11	674.31	.4711
High Proficiency	974.56	310.15	.3181

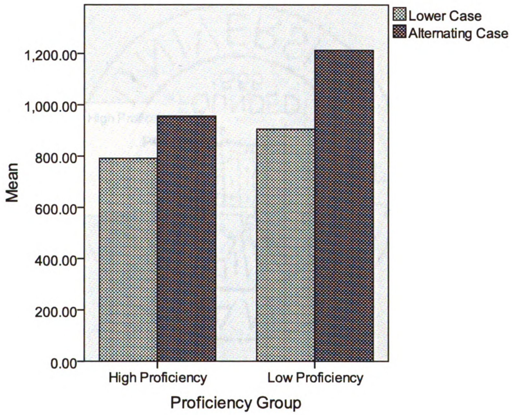


Figure 7. Means of lower case and alternating case reaction times by proficiency group.

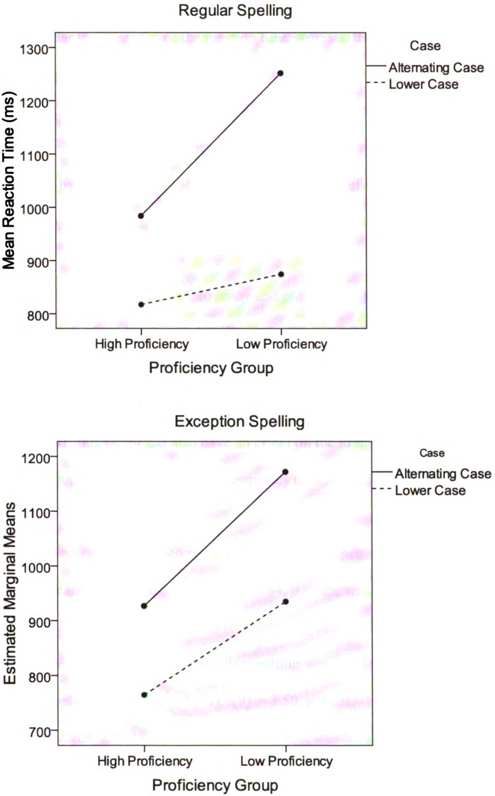


Figure 8. Interaction between case and proficiency within spelling regularity conditions.

Accuracy

Similar to the results of the reaction time analysis, there was a main effect of proficiency group when participants' error rates were analyzed both by participant, $F(1,18) = 63.63, p < .001$, and by item, $F(1,36) = 27.58, p < .001$. In addition, there was a main effect of case when error rates were analyzed by participant $F(1,18) = 10.45$, $p < .01$, and by item $F(1,36) = 6.75, p < .05$. The main effect of frequency was also significant by both participant, $F(1,18) = 93.89, p < .001$, and by item, $F(1,36) = 46.44, p < .001$.

The interaction between frequency and proficiency group was also significant, by participant: $F(1,18) = 37.82, p < .001$, by item: $F(1,36) = 21.81, p < .001$. Post-hoc paired samples t-tests showed that the lower proficiency group showed significantly different error rates between high (2.89%) and low frequency words (31.70%), $t(9) = -6.87, p < .001$; this factor did also have an effect on the high proficiency group (1.50% for high frequency and 8.03% for low frequency), although this difference did not reach significance (Figure 9).

The interaction between proficiency group and spelling regularity was also found to be significant, by participant: $F(1,18) = 4.84, p < .05$, and by item: $F(1,36) = 4.14, p < .05$. Post-hoc paired-sample t-tests revealed a significant difference between the high proficiency readers' error rate for regular spelling words ($M = 7.5\%$) and that for exception spelling words ($M = 2.0\%$), $t(9) = 3.50, p < .01$. This difference, $M = 16.2\%$ for regular words, $M = 18.4\%$ for exception words, was not significant in the case of the low proficiency readers, $t(9) = -.756, p > .4$.

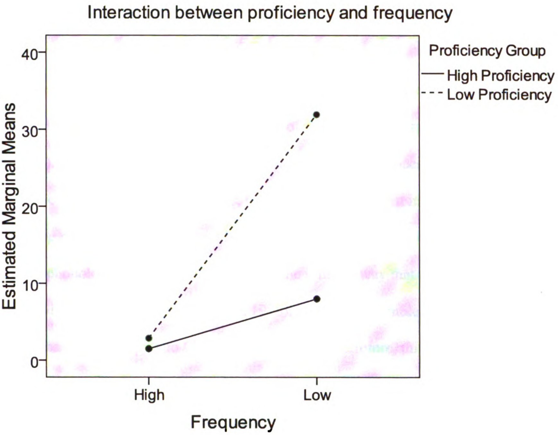


Figure 9. Error rate interaction between proficiency group and frequency.

The interaction between proficiency group and case was significant in the item analysis, $F(1,36) = 4.51, p < .05$; in the participant analysis, however, it approached but did not reach significance, $F(1,18) = 4.20, p = .055$. In addition, none of the three-way or four-way interactions were significant (all $F(1,18) < 1.38$, $p > .2$ by participant, $F(1,36) < 0.84$, $p > .3$ by item).

Discussion

The lexical decision task was designed to measure the effect of proficiency on readers' ability to process the constituent letters of words (i.e., in response to Research Question 1). The processing of individual letters in words is a process that does not exist in the native language of the participants, Chinese, but is prevalent in English. In addition, the processing of constituent letters in words can be seen as a measure of assembled phonology, which is theorized to be the primary way that words are recognized in English (Lesch & Pollatsek, 1998); however, assembled phonology is also a concept that does not apply to Chinese orthography. Therefore, this experiment sought to determine whether the processing of constituent letters and assembled phonology while reading L2 English are developmental; that is, whether they change along with L2 proficiency.

Based on the literature, it was hypothesized that L2 proficiency would have an effect on L2 English word recognition in terms of processing the constituent letters of words. Overall, the results of this experiment support this hypothesis, thus answering affirmatively to the first research question. Both the reaction time results as well as the accuracy results showed that alternating case text had a significantly greater effect on the low proficiency readers than on the high proficiency readers. That is, the high proficiency readers showed less interference from loss of visual word shape information and the processing of the individual letters of words. In the item analysis of the reaction time data, however, this effect was constrained to the words that were regularly spelled (research question 1b), and did not appear in the words that had exception spelling. In the analysis of all alternating case stimuli as well as regular spelling alternating case stimuli, restructuring of cognitive processes, rather than mere speed-up, was shown. Through the analysis of the reaction time data, SDrt, and CVrt, it was shown that CVrt decreased significantly, indicating restructuring.

The interaction between proficiency, case, and spelling regularity may be consistent with the theory that phonology was activated while the alternating case words were parsed. In regular spelling words, grapheme-phoneme correspondence rules are more readily accessible to the reader. If an effect of case were found across both regular and exception word conditions, then it could be said that readers were not affected by these differences in the relationships between the letters and their sounds, since spelling regularity is a measure of the relationship between letters and their sounds. This three-way interaction was, however, only found in the item analysis of the reaction time data, and in none of the other analyses. In all other analyses of the reaction time data, there was a two-way interaction between proficiency and case, without influence from spelling regularity.

The absence of a three-way interaction between proficiency, case, and frequency shows that the effect of proficiency on reading constituent letters of words did not depend on the frequency of the words. In answer to research question 1a, frequency did not have an effect on the relationship between case and proficiency group.

These findings suggest that word recognition processes are, in fact, restructured as L2 readers gain in L2 proficiency, perhaps in a number of ways. The results of this experiment may provide an extension to previous research regarding intra-word processing/decoding (e.g., Brown & Haynes, 1985; Holm & Dodd, 1995; Koda, 1999; Wang, Koda, & Perfetti, 2003) to show that intra-word processing can develop together with proficiency. In addition, assuming that the processing of the constituent letters of a word does lead to assembled phonology during word recognition as theorized, these findings may also demonstrate that the degree of phonological activation during visual word recognition may be developmental as well.

In addition, the significant interaction between proficiency and spelling regularity found in the error rate analysis and the three-way interaction between proficiency group, spelling regularity, and case found in the reaction time analysis may also have relevance to the issue of activation of phonology while reading. The low proficiency group showed slower reaction times for the exception words than the regular words (as was expected based on the literature) and no significant difference in accuracy between the regular spelling words and the exception spelling words (found in the error rate analysis), suggesting a lack of sensitivity to spelling-sound relationships. In effect, the low proficiency readers were either reading words the same, no matter their differences in grapheme-phoneme correspondence regularity, or were possibly affected by the compounding of weaker intra-word processing abilities with exception spelling of words. The fact that the high proficiency readers took longer to read words with regular GPC rules and made a significantly larger number of errors on these words was unexpected, as it is typically exception words on which readers make more errors. However, this result may be evidence of a sensitivity to grapheme-phoneme correspondence rules that the high proficiency readers have developed, and which the low proficiency readers do not have. It is possible that because of their larger vocabulary and greater exposure to printed language, high proficiency readers have identified which words need to be processed in an addressed manner, and which words can be processed in an assembled manner; i.e., high proficiency readers processed exception words in an addressed manner because of the inconsistent GPC rules contained within them, while they processed words with regular GPC rules in an assembled manner. Assembled processing typically takes longer than addressed processing (Fender, 2003), which would explain why the assembled (regular spelling) words had longer reaction times for the high proficiency readers.

In addition, other robust main effects and interactions are also worth noting. In all analyses, there were significant main effects of proficiency group. This shows that, in general, high proficiency readers read words more quickly and more accurately than low proficiency readers did. As was predicted, word frequency also had a large effect on both groups. However, both the reaction time analyses and the error rate analyses showed that frequency had an even greater effect on the low-proficiency group than the high proficiency group. This is to be expected, as higher-proficiency readers would tend to have larger vocabularies, which would typically include greater exposure to a larger number of low-frequency words, than lower-proficiency readers would.

Case also had a major effect on both groups. Even though alternating case had a significantly greater effect on the low proficiency readers than the high proficiency readers, the effect on the high-proficiency readers was significant as well. This may show that although readers' processing of the constituent letters improves with proficiency, there is still an L1 effect even at higher L2 proficiency levels. This effect could, however, also be a task effect in that it can be assumed that even native English speakers would slow down to some extent when reading text in alternating case versus all lower case, no matter their ability to process the constituent letters of words. Text in alternating case introduces an added effect in that letters are encountered in places where they are not expected to be (that is, capital letters are in the middle or at the end of words). Thus, the 'translation' of the capital letters into their lower case counterparts may also introduce a certain degree of latency that is unavoidable, and most likely would occur with native speakers of English as well.

Although this measure of the processing of constituent letters of words gives a good approximation of the activation of the phonology of words, it does so in a somewhat indirect way. In addition, lexical decision tasks may involve a minimal amount of semantic access, as it may be possible to make a decision as to a string of letters' status as a word or not without activating the meaning of the word. For that reason, a second experiment was conducted that measures reliance on phonological information in a more direct manner and involves explicit access to meanings of words, although without the ability to incorporate the word factors investigated in the lexical decision task.

Experiment 2

Experiment 2 addresses the second research question: "When high L2 proficiency readers read English words for meaning, do they show more reliance on phonological information compared to low proficiency readers? Do lower L2 proficiency readers show heavier reliance on orthographic information than high L2 proficiency readers?"

Experiment 1 used text in alternating case to investigate the effect of visual familiarity; as in previous studies using the same paradigm, these tasks were designed based on the assumption that visually-unfamiliar words (alternating-case words, in the above experiments) would be processed in an assembled manner, which would activate the phonology of the constituent letters. However, lexical decision tasks do not provide a way to compare relative reliance on visual and phonological processing, but instead force the reader to process the constituent letters of words. There may be a difference between a reader's *ability* to read a certain way when forced, and their tendencies when they *actually* read. In addition, Experiment 1 may show changes in how readers process the constituent letters of words, but there is the possibility that high-proficiency readers may have been less affected by the alternating case because of greater familiarity with the letters themselves. Therefore, a third task was also conducted that attempted to gauge how much participants relied on phonology during word recognition without forcing them to process words in a specific way. This task is based on a task developed by Van Orden in 1987 for use with L1 English speakers, and used subsequently by Wydell, Patterson, & Humphreys (1993) with L1 Japanese, and by Wang, Koda, and Perfetti (2003) in an L2 (English) context with L1 Chinese and L1 Korean. This experiment is a partial replication of a part of Wang et al. (2003) using the same stimuli, but with different populations (and therefore a different between-subjects independent variable).

Participants

The same extreme groups of participants that were included in Experiment 1 were included in Experiment 2 as well.

Materials

The stimuli from Wang, Koda, and Perfetti (2003) were used, with permission, in this experiment. The two contrasts that were included in the stimuli were phonological similarity and spelling similarity (Table 6). Phonological similarity had two levels: phonologically similar (homophonic) foils of category exemplar words, and orthographically similar controls. Spelling similarity also had two levels: similarly spelled, and less similarly spelled. That is, half of the phonological foils and orthographic controls were spelled similarly to the exemplar word that was being primed for, while the other phonological foils and orthographic controls were words with spellings that were less similar to the word that had been primed for. Therefore, in total, there were four experimental conditions: 1) similarly spelled phonological foils (e.g. "feat" for the category "a body part"), 2) similarly spelled orthographic controls (e.g. "feel"), 3) less similarly spelled phonological foils (e.g. "rows" for the category "a flower"), and 4) less similarly spelled orthographic controls (e.g. "robe"). Stimuli consisted of 9 category words or phrases for similarly spelled words, and 8 categories for the less similarly spelled words. Categorizations of "more" or "less" similarly spelled were made by Wang et al. (2003). All stimuli are listed in Appendix F.

When the stimuli were developed by Wang et al. (2003), the stimuli were also judged by ESL instructors as to the words' familiarity to ESL learners. This was because of effects that have been found in terms of frequency in previous studies. Wang et al. chose phonological foils and orthographic controls of moderate familiarity on the basis that more familiar words would be so-called "sight words", which would typically be recognized with a lesser degree of phonological activation, and less familiar words would be more likely to be recognized with a higher degree of phonological activation. It can be assumed that the familiarity judgments of the ESL instructors in that study, if accurate for the participants in Wang et al. (2003), would most likely also be accurate for the population included in the present study.

In total, all participants completed 103 trials, of which 51 were critical trials and 52 were distractors. Distractors were included so that participants would not guess the purpose of the study because a high number of stimuli were similar to actual category words. Of the 51 critical trials, 17 were category exemplar words, 17 were homophones of the exemplar words, and 17 were words that were orthographically similar to the exemplar words. Of the 52 distractors, 34 were "Yes" trials (i.e. words that did fit the category) and 18 were "No" trials. As in Wang et al. (2003), this was in order to offset the larger number of "No" trials among the critical trials. Participants completed 12 practice items before the real test began, receiving feedback on the first two. This task was also presented using the DMDX software package (Forster & Forster, 2003).

Procedure

Stimuli were presented on a Windows XP-based computer in black text on a white background in 24-point font; participants were seated approximately 50cm from the monitor with blinder walls on either side. Instructions are in Appendix G.

Table 6

Examples of phonological and orthographic similarity

	Category Prime	Phonological Foil	Orthographic Control
Similar	A body part	feet	feat
Less Similar	Part of a person's face	knows	snobs

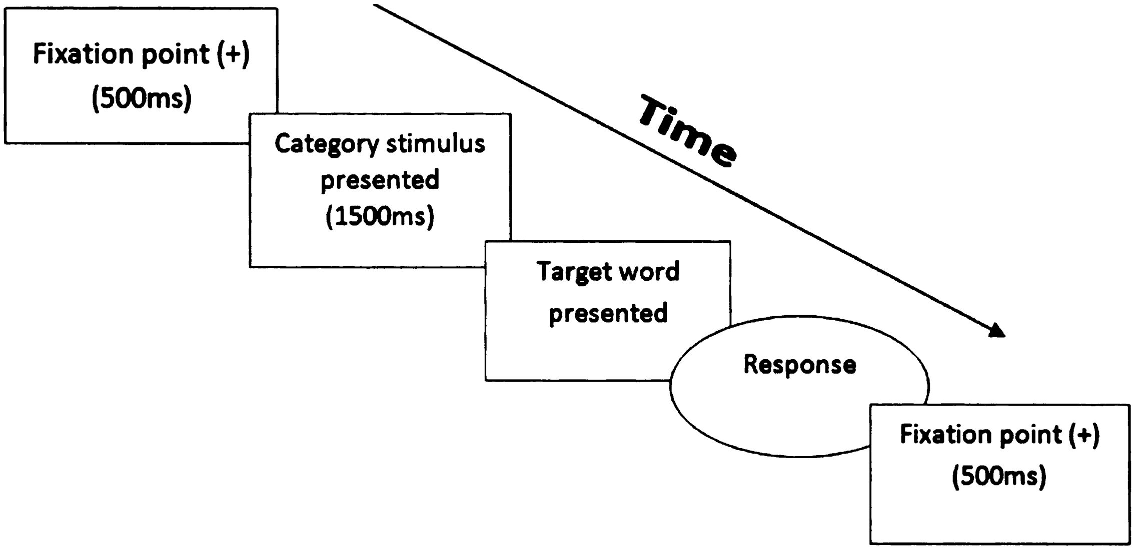


Figure 10. Paradigm used in Experiment 2.

Participants were first presented with a fixation point (+) for 500ms, after which they were shown a category stimulus for 1500ms. Next, the category prime disappeared, and a target word (or a distractor) appeared. Participants then decided whether the target word (or distractor) fit the category or not by pressing the red (No) or green (Yes) keys (described above). After making a decision, participants were again presented with a fixation point (+) for 500ms before presentation of the next category prime. Figure 10 shows a flowchart of the procedure. All times are consistent with Wang et al. (2003).

It is theorized that if a participant makes more mistakes (or has higher reaction times) on the phonological foils than on the orthographic controls, then the reader is relying substantially on phonology while reading for meaning. If, on the other hand, a participant makes more false-positives on more similarly spelled foils than less similarly spelled foils, then the reader is relying on orthography.

Results

The experiment implemented a 2 (proficiency group) x 2 (phonological similarity) x 2 (spelling similarity) repeated measures factorial design, with proficiency group as a between-subject independent variable and phonological similarity and spelling similarity within-subject independent variables in the participant analysis. In the item analysis, the word factors were between-item independent variables, and proficiency was a within-items independent variable. Dependent variables in both the participant and item analyses were the error rates and reaction times of the category judgments. Analyses were performed using repeated measures factorial ANOVAs.

Accuracy

Table 7 shows the means and standard deviations of the error rates for each of the four conditions for both proficiency levels. The analysis showed a main effect of proficiency group, reliable by participant, $F(1,18) = 18.20, p < .001$, and by item. $F(1,30) = 42.24, p < .001$. Overall, high proficiency readers made fewer errors than low-proficiency readers in all four experimental conditions (Figure 11).
In addition, there was a main effect of phonological similarity that was reliable by participant: $F(1,18) = 10.02, p < .01$, indicating that, overall, more errors were made on phonological foils than on orthographic controls. This was, however, not significant in the item analysis, $F(1,30) = 2.49, p > .1$. The main effect of spelling similarity was not significant by participant or by item.

More pertinent to the present study, however, were the interactions between the word-based variables (phonological and orthographic similarity) and proficiency. Although neither two-way interaction between phonological or orthographic similarity and proficiency were significant, there was a significant three-way interaction, both by participant, $F(1,18) = 15.94, p < .001$, and by item, $F(1,30) = 5.97, p < .05$. Post-hoc paired sample *t*-tests revealed a significant effect by participant of phonological similarity for the high proficiency group; that is, the high proficiency group had significantly higher error rates on the phonological foils, showing a sensitivity to the phonology of the words, $t(19) = 2.77, p < .05$; however, this was not reliable by item $t(16) = 1.56, p > .1$. Phonological similarity did not have an effect on the low proficiency group, by participant, $t(19) = 1.03, p > .3$, or by item, $t(16) = 0.72, p > .4$. Post-hoc paired sample *t*-tests on the effect of spelling similarity did not show any significant differences for either high or low proficiency readers, by either participant or item.

Table 7

Means and standard deviations of error rates for each condition, in percents

	Similarly Spelled		Less Similarly Spelled	
	M	*SD*	*M*	*SD*
High Proficiency				
Phonological Foils	18.88	19.65	13.75	12.43
Orthographic Controls	7.77	9.14	6.25	10.62
Low Proficiency				
Phonological Foils	32.21	26.43	51.25	19.94
Orthographic Controls	45.53	19.22	26.25	18.11

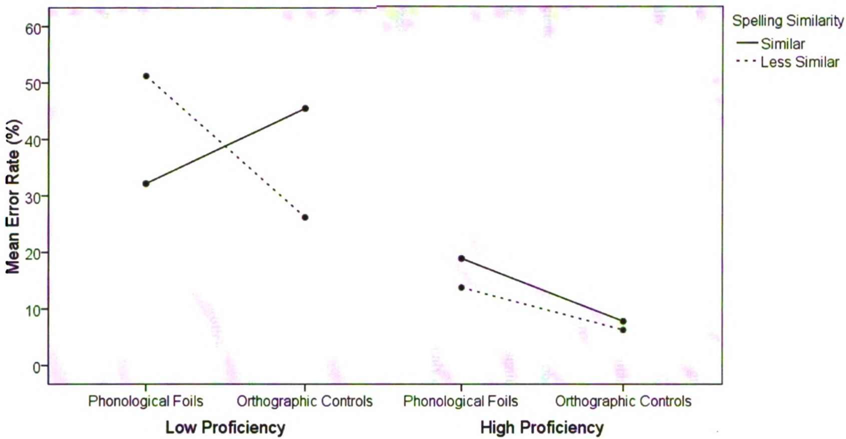


Figure 11. Category judgment accuracies in the four conditions by proficiency group.

In addition, an unexpected result appeared within the similarly spelled conditions (Figure 12). Whereas the high proficiency group made more false-positives on the phonological foils, as was expected as well as what was shown across all conditions by the Chinese group in Wang et al. (2003), the low proficiency group in the present study actually made more false-positives on the orthographic controls within the similarly spelled conditions. This result was not found in the less similarly spelled stimuli, however, where low proficiency readers made more errors on phonological foils. High proficiency readers, on the other hand, made more errors on phonological foils than orthographic controls regardless of differences in spelling similarity.

Reaction Times

As is customary in analyses of reaction time data, only correct responses were analyzed. Reaction time data were trimmed to two standard deviations from the mean of each participant. As in the error rate analysis, a 2x2x2 repeated measures factorial ANOVA was performed.

The analysis of the reaction time data showed a main effect of proficiency group, which was significant by participant, $F(1,18) = 11.64, p < .01$, and by item, $F(1,30) = 73.11, p < .001$. This shows that, overall, the high proficiency readers made faster decision times than the low proficiency readers. There was a significant interaction between proficiency group and phonological similarity in the item analysis, $F(1,30) = 7.63, p < .01$. Although the high proficiency readers' mean reaction time only increased from 1157.18ms ($SD = 187.69$) on the phonological foils to 1167.06ms ($SD = 163.62$) on the orthographic controls (an increase of less than 10ms), the low proficiency readers' mean reaction time increased from 1477.71ms (SD = 289.247) on the phonological foils to 1788.18ms (SD = 396.94) on the orthographic foils, an increase of over 300ms (Figure 13). Although this difference was significant by item, it was not significant by participant, $F(1,18) = 2.96, p > .1$.

None of the interactions between spelling similarity and proficiency group, phonological similarity and spelling similarity, or the three-way interaction between these were significant when analyzed by either participant or item (all $F(1,18) < 2.96, p > .1$).

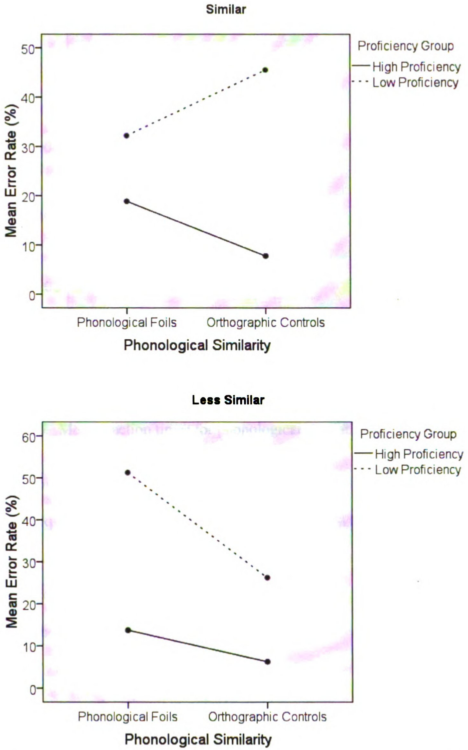


Figure 12. Error rates of phonological similarity in the similarly-spelled and less similarly spelled conditions.

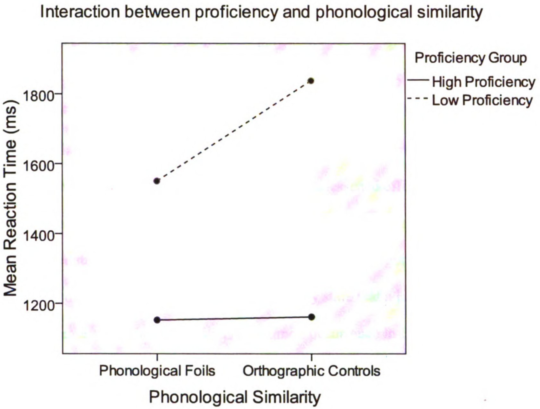


Figure 13. Mean reaction times for phonological similarity by proficiency group.

Discussion

Experiment 2 set out to answer the questions, "When high L2 proficiency readers read English words for meaning, do they show more reliance on phonological information compared to low proficiency readers?" and, "Do lower L2 proficiency readers show heavier reliance on orthographic information than high L2 proficiency readers?" This experiment used differences in phonological similarity and spelling similarity among the stimuli in order to provide answers to these questions. Through identical methodology, Wang et al. (2003) showed that L1 Chinese speakers relied less on phonological information than L2 English readers whose L1 orthography is also alphabetic (Korean). In addition, that study showed that L1 Chinese speakers showed a greater sensitivity to orthographic similarity when reading L2 English than alphabetic (Korean) readers. The present study investigated the effects of these factors within a group of L1 Chinese readers of L2 English, but with differing levels of proficiency. It was hypothesized that processing strategies transferred from the L1, those demonstrated by Wang et al. (2003), would be more evident in low proficiency readers, while high proficiency readers would show patterns that reflect a lessened effect of L1 orthography. In the analysis of the results, the error rate data is likely to be more useful than the reaction time data. Wang, Koda, and Perfetti (2003) had few significant results from the reaction time data, similar to the present study. They state that a lack of significant interactions among reaction time data is a common occurrence in L2 research because of low L2 accuracy levels (Juffs, 2001). In the present experiment, as well, there could have been an effect of the very high error rates that occurred due to the paradigm of the experiment, which induced errors through the use of words that were homophones of, or spelled similarly to, words that had been primed for. Because the reaction time analyses included only correct responses, in some cases there could have been a only small number of accurate results remaining after eliminating incorrect responses, which may not provide adequate information regarding the processes that occur in the populations involved. Therefore, this limitation should be considered when interpreting any results (significant or non-significant) from the analysis of the reaction time data.

Concerning the first part of research question two, whether high proficiency readers show more reliance on phonological information compared to low proficiency readers, the results suggest that the answer is in the affirmative. For the high proficiency readers, phonological similarity to primed words (i.e. homophones) caused readers to make more errors than when words were orthographically similar. On the other hand, such an effect was only evident in low proficiency readers under certain orthography-based conditions (orthographically less-similar words). Therefore, it appears high proficiency readers had a more general reliance on phonology than low proficiency readers did.

Regarding the second part of research question two, whether low proficiency readers show heavier reliance on orthographic information than high proficiency readers do, the results also indicate a positive answer. The low proficiency readers showed much more variation as a result of spelling similarity than the high proficiency readers did, even to the extent that it changed their reliance on phonology. The high proficiency readers showed nearly identical patterns when reading similarly spelled stimuli and less similarly spelled stimuli, thus showing a lesser degree of reliance on this orthographic information. The low proficiency readers, on the other hand, showed reliance on phonological information when interference from orthographic similarity was less, and reliance on orthographic information when orthographic similarity was greater.

However, this is not to say that the low proficiency readers only showed reliance on orthographical information, without reliance on phonological information; rather, the low proficiency readers showed a very high degree of reliance on phonology within the less similarly spelled conditions. That is, only when the influence of orthography was lessened, as was the case in the less similarly spelled stimuli, did the effect of phonological similarity appear. This could indicate that although low L2 English proficiency readers whose L1 is Chinese do rely on phonological information to an extent while reading for meaning, they rely primarily on orthographic information, and secondarily on phonological information.

Overall, these results suggest that low proficiency L2 English readers whose L1 is Chinese show less reliance on phonological information, and greater reliance on orthographic information, than high proficiency L2 English readers. The high proficiency readers showed patterns that are nearly identical to those shown by the alphabetic readers in Wang et al. (2003), suggesting that as proficiency increases, phonological activation may become more prevalent, and therefore a larger factor, during word recognition. The low proficiency readers, on the other hand, showed results that were somewhat different from those shown by the Chinese group in Wang et al. (2003). In Wang et al., the participants were 15 out of 20 participants were recruited from an ESL program, although from the intermediate and advanced level courses, whereas the low proficiency readers in the present study were recruited from lower level courses. The remaining 5 L1 Chinese participants in that study were first year graduate students, which would seem to be more similar to the high proficiency group in the present study. However, the mean MTELP score for the L1 Chinese group in Wang et al. was 62.13% (for all three sections of the MTELP), whereas the high proficiency group's mean MTELP score in the present study was 83.67%, while the low proficiency group's mean score was 37.67% (for the grammar and reading sections of the MTELP). Therefore, the participants in Wang et al. could be seen as between the high and low proficiency groups in the present study. In Wang et al., however, orthographic information had a greater impact on the L1 Chinese group than did phonological information. In the present study, however, both orthographic information and phonological information had large impacts on word recognition by the low proficiency readers, suggesting that low proficiency readers do have a certain level of reliance on phonological information as well.

General Discussion

The present study sought to investigate differences in word recognition processes that arise from differences in L2 proficiency. Overall, the results of the two experiments suggest that with increased L2 English proficiency, intra-word processing of constituent letters of words increases, as does reliance on phonological information. In addition, the results suggest that with proficiency, the cognitive processes of L2 English learners whose L1 is non-alphabetic are restructured, rather than merely sped-up. These results are consistent with results shown by Chikamatsu (2006), which investigated L2 Japanese reading processes of L1 English speakers (i.e., alphabetic L1 and morphographic L2, opposite of the present study) and thus restructuring of processes in order to fit (morphographic) Japanese orthography (i.e., a *greater* reliance on orthographic information by the high proficiency group). The results of the present study are also consistent with predictions made, although not investigated, by Wang, Koda, and Perfetti (2003).

Taken together, the results of the two experiments suggest that the effects of L1 orthography on word recognition processes (i.e. lesser reliance on, and lesser activation of, phonology) might diminish as L2 proficiency increases. This was seen in the present study through the use of two very different methodologies in order to reduce any effects that may be resultant from the type of task used. Experiment 1 investigated these phenomena in terms of processing constituent letters of words and Experiment 2 did so in terms of reliance on phonological information.

In Experiment 1, it was shown that high L2 English proficiency readers showed faster intra-word processing with fewer errors as well as greater sensitivity to grapheme-phoneme correspondence than did low L2 English proficiency readers. High proficiency readers processed the individual letters of words, theorized to activate the phonology of those letters, more quickly and with a higher accuracy rate than lower proficiency readers. An analysis of the differences between the coefficients of variance of the participants in both proficiency groups suggest that the differences in intra-word processing between high and low proficiency readers may be due to the restructuring of cognitive processes, rather than simple speed-up.

In Experiment 2, high L2 English proficiency readers relied more on phonological information than did low L2 English proficiency readers. In addition, it was shown that low L2 English proficiency readers may be more sensitive to orthography than high proficiency readers. These results would seem to support the results of Experiment 1, showing a decrease in L1 orthographic transfer in higher L2 proficiency readers.

These results may also be relevant to the application of the orthographic depth hypothesis to L2 reading. The ODH hypothesizes degrees of phonological reliance due to orthography, and previous cross-orthographic word recognition research has applied it in this way, without regard to other factors that may affect word recognition. Previous research has applied the ODH to cross-orthographic word recognition in terms that are absolute: that is, that the effect of the L1 orthography on L2 reading is constant and unvarying. The results of the present study, however, suggest that the degree to which L1 orthography affects L2 reading may actually be more variable, with higher proficiency readers showing less of an L1 orthographic effect than lower proficiency readers. Future studies of L2 word recognition must pay closer attention to L2 proficiency as a potential source of variation among learners.

These results also highlight differences between L1 and L2 literacy development. When learning to read, native English-speaking children go through phases of literacy development that include the amalgamation of constituent letters of words at the lower end of development, and addressed recognition of entire words (i.e., the reading of "sight words") at the higher end of development (Ehri, 1995). In fact, for native English-speaking children, the ability to process the constituent letters of words, and the ability to phonologically recode those letters, is a requisite skill of sight word reading (Ehri, 1998). The present study, however, suggests that it may not necessarily be this way in the development of L2 English literacy, and in fact, depending on the orthography of the reader's L1, these skills may develop in an order opposite of L1 readers. Similar to L1 reading acquisition, however, is that highly proficient readers have abilities to process words both with reliance on phonological information as well as reliance on orthographic information.

Limitations and Future Research

While the present study showed what appear to be consistent results triangulated through the use of independent experimental paradigms, they did so with a very limited sample size ($N = 20$). Future research investigating developmental changes in the influence of L1 orthography on L2 word recognition should be performed incorporating larger numbers of participants in order to find results that are more generalizable.

In addition, as this study was concerned with the effect of L2 proficiency on L2 word recognition, a cross-section of proficiency levels was used. Ideally, a longitudinal study should be performed in order to assess changes as a result of L2 proficiency within individual learners.

Furthermore, the present study only investigated recognition of single words, which is not how people usually read. Previous research (e.g. Chikamatsu, 2006; Fender, 2003) has shown that L2 learners may recognize words differently based on the context in which the words appear. It would be beneficial for future studies to investigate effects of L2 proficiency on not only single words, but also words within a text (sentence, paragraph, etc.) as well.

The present study assumes, based on the orthographic depth hypothesis, that differences between the low proficiency readers and the high proficiency readers were in response to the alphabetic nature of the orthography of the L2. It would be prudent for future research to incorporate other alphabetic L2s as well in order to eliminate any possible effects that are specific to English, and to find which effects can be generalizable to alphabetic orthographies as a whole. It would also be informative to compare results of L2 English readers with L1 English readers to see if word recognition processes of high proficiency L2 English readers become more similar to those of L1 English readers.

In the same vein, it would also be beneficial for future studies to include L2 English learners of not only a variety of L2 proficiency levels, but L1s as well. It is possible that the developmental results found in this study are common to all learners of English, regardless of the nature of the L1 orthography. For this reason, high and low L2 proficiency readers of a greater number and variety of L1s should be included in future studies. In addition, it might be the case that high proficiency L2 English readers of any type of L1 will read words in similar ways; that is, with increased L2 proficiency, L2 readers may converge on a specific type of processing that is optimal for the L2. This is suggested by very close similarities between the high proficiency group in the present study (who have a morphographic L1 orthography) and the alphabetic L1 group (Korean) in Wang et al. (2003).

In terms of pedagogical applications, if the results found in the present study are found to be generalizable, future pedagogical studies of learner training that encourages processing of constituent letters in words and reliance on phonology might be also useful (a vein of research already begun by Akamatsu, 2007). If processing of constituent letters of words and reliance on phonological information are, in fact, traits of high proficiency readers, it may be useful to investigate whether these traits are an extraneous result of increased proficiency only, or whether the acquisition of these traits can produce higher-level readers.

Appendix A

You are invited to participate in a research study about reading in a second language.

The study will investigate how people whose native language is Chinese read English. You are able to participate in this study if:

1) Your native language is Chinese, and
2) You are between 18 and 40 years old.

Procedure:
The study will involve attending 1 session, arranged with Ryan Miller. During this session, you will:
1) fill out a short questionnaire about your language learning experiences,
2) take a test to measure your English language proficiency, and
3) read words and sentences on a computer screen and push buttons on the keyboard.

Session times:
<Session times, dates, locations>
Or, you can schedule a time with Ryan Miller by emailing ryantm@msu.edu.

Voluntary participation:
Your participation in this study is completely voluntary. You may refuse to answer any questions, and you may withdraw from the study at any time.

Compensation:
For your participation in this study, you will receive $15 cash payment.

Risks:
Your participation in this study will not affect your enrollment in the university, and it will not affect your grades in your classes.

Benefits:
Your participation in this study will give you some extra practice in reading English. In addition, it contributes to second language research.

Confidentiality:
You will not be required to give your name or any other personally identifiable information. You will be identified only by a number throughout the data collection process as well as any further publication.

Contact information:
If you have any questions or concerns, contact Ryan Miller at ryantm@msu.edu, by phone at 517-355-0800, or by mail at A-714 Wells Hall, Michigan State University, East Lansing, MI, 48824

Appendix B

Background Information

Subject ID: ______

1) Gender (circle one): Male Female
2) Age: _______ years old
3) Country of birth: ___________________
4) What is your native language? Chinese English Other (Language:____________)
5) If your native language is Chinese, did you use pinyin (拼音) when learning Chinese?

   Yes No

   a. If yes, how old were you when you started using pinyin (拼音)? _____ years old
6) What language(s) do you speak at home now? (circle)

   Chinese English Other (Language:___________)

   a. If more than 1, what percent of the time do you speak each:

      Chinese: ________% English: ________% Other: ________%

7) In what languages did you receive your education before university?

   Chinese English Other (What language? _____________)

   a. If more than one, please give the approximate number of years for each language

      Chinese: ________ English: ________ Other________

8) How many years and months have you lived in Michigan? ______ years ______ months

9) Have you ever lived in another English-speaking place besides Michigan?

   Yes No

a. If yes, where, and for how long? Place: ________, ___years ___ months

Place: ________, ___years ___ months

10) Please mark what you think your ability is in each of the following.

Use the scale: 1 = Poor, 2 = Fair, 3 = Good, 4 = Very good, 5 = Native/Native-like

	Listening	Speaking	Reading	Writing	Grammar
Chinese					
English					
Other (Language:__________)					

11) How long do you usually spend **each week** doing the following activities? How long **did** you spend before **last** semester? Please write the number of **hours** and **minutes**.

	NOW	BEFORE **last** semester
Reading in English	____hours____minutes	____hours____minutes
Writing in English	____hours____minutes	____hours____minutes
Listening in English	____hours____minutes	____hours____minutes
Speaking in English	____hours____minutes	____hours____minutes
Reading in Chinese	____hours____minutes	____hours____minutes
Writing in Chinese	____hours____minutes	____hours____minutes

12) What do you usually read in English, and how long do you spend reading them **each week**? If 'other', please write what you read and for how long.

Newspapers:___hours___minutes Other (__________): ___hours____minutes

Magazines: ___hours___minutes Other (__________): ___hours____minutes

Novels: ___hours___minutes Other (__________): ___hours____minutes

Web sites: ___hours___minutes Other (__________): ___hours____minutes

Appendix C

Debriefing

The study you have participated in concerned impact of second language proficiency on the process of recognizing words while reading. Specifically, it measured how native Chinese speakers use sound while reading in English. One task that you performed measured the time it took you to recognize single words. Some words were more common words than other words. This is because the brain tends to use the sounds of words less for common words than for uncommon words. Also, some words had spellings that follow common spelling patterns and some were more unusual. This is because the brain may store words that have less common spelling patterns as a larger piece, rather than as a combination of sounds. Some words were not real words. These words were included to test your ability to recognize real words as opposed to unreal words. Some words were in lower case (lower case) and some words were in alternating case (aLtErNaTiNg CaSe) because when you read words in alternating case, your brain cannot recognize the words by their shape, and instead your brain must figure out the sounds of the words. Another task was similar to the first task, but had the words in sentences. This was to measure changes that might result from reading to understand meanings of words and sentences. The remaining task was to measure how you read words for meaning, but not in sentences.

As compensation for my participation in the research study, I acknowledge receipt of $15 cash from ______________.

(researcher name)

____

Signature

Appendix D

High Frequency – Regular Spelling		**High Frequency – Exception Spelling**	
such	state	good	work
men	still	house	again
place	fact	word	does
come	want	come	water
home	most	took	once
gave	week	head	knew
tell	free	group	young
help	class	half	girl
soon	road	says	love
turn	main	dead	move
Low Frequency – Regular Spelling		**Low Frequency – Exception Spelling**	
pitch	neat	bowl	sweat
beam	fluid	naive	flood
sank	rage	earn	urge
ripe	wax	blew	fur
tower	dot	knock	log
pump	trunk	wool	glue
slip	deed	knit	aisle
float	weld	sew	swamp
slam	peel	worm	swarm
spark	fade	ivory	dial

Non-word Distractors				
balf	dost	gial	maig	pome
bant	dramp	gorm	mand	poom
boad	dreat	gream	mave	poot
bood	drep	haide	meafe	reat
bool	fep	hepe	mide	rilp
boose	flace	hupe	mif	rone
bork	flape	jart	mive	roshy
bove	foll	kolm	nad	rull
cipy	forli	kouf	nar	starm
clow	foth	lairm	natch	stook
dase	fove	lare	nouse	sut
deak	frake	lear	nowl	tays
deght	fregh	lere	orm	tew
dit	frep	lipht	phint	trom
dolp	frug	lity	plue	tross
dool	fushe	lor	poes	vank

Appendix E

In this task, you will see some text on the screen. You need to decide if the text is a real English word or not. Some words will be real words, and some will not be real words.
Press the SPACEBAR to continue.

If you think the text is a real word in English, press the GREEN button as soon as you recognize the word. If you think the text is not a real word in English, press the RED button.
Press the SPACEBAR to continue.

Try to answer as quickly and accurately as you can, but be sure to read each word before you press a button.
Press the SPACEBAR to continue.

Some words will be written in a mixture of UPPERCASE letters and lowercase letters. This means that 'dog' is still a word if it is written 'DoG' or 'dog'. For both of these, you would press the GREEN button (meaning Yes, it is a word).
Press the SPACEBAR to continue.

First, you will have 10 practice questions. Remember, push the green button if what you see is a word in English, and press the red button if you think it is not a word in English.
Push the SPACEBAR to continue.

Press the SPACEBAR to go on to the 10 practice questions.

Ready? Here we go...

[Practice items]
big
gorp
aPpLe
cup
pHoNe
wOfP
cidno
wIrE
fLoMp
klim

Appendix F

Category name	Exemplar	Homophone	Spelling Control
	Similarly spelled foils		
A feature of an ocean shore	beach	beech	bench
Part of a sandwich	bread	bred	bead
A small stream	creek	creak	cheek
A body part	feet	feat	fees
Transportation in the sky	plane	plain	plans
Type of weather	rain	rein	ruin
Used to get up or down	stair	stare	stars
The end of your feet	toe	tow	toy
An alcoholic drink	wine	whine	wink
	Less similarly spelled foils		
An animal	bear	bare	beat
A breakfast food	cereal	serial	several
A part of a plant	flower	flour	flowed
Part of a person's face	nose	knows	snobs
A passage used by vehicles	road	rode	rods
A flower	rose	rows	robs
Part of a boat	sail	sale	salt
A sea animal	whale	wale	wheel

Distractors			
Category	Stimulus	Category	Stimulus
Used for communication	phone	A hot drink	tea
Something you can read	book	A place on campus	union
A place to live	house	A place on campus	stadium
Something to hold liquids	cup	A sport	soccer
A person	friend	A sport	tennis
Something you carry	bag	A thing you wear on your foot	shoe
Something you wear	coat	A thing you wear on your foot	sandal
A fruit	lemon	A place you work	office
A fruit	apple	A fruit	book
A kind of furniture	chair	Something you wear	plane
A kind of furniture	table	A vehicle	cloud
A place in a house	kitchen	A food	book
A place in a house	bedroom	A person	table
Something you watch	TV	An animal	chair
A type of clothing	shirt	A computer part	friend
A type of clothing	hat	A place in a house	apple
A color	blue	A hot drink	light
A color	green	An animal	wire
An animal	mouse	A type of weather	pig
An animal	rabbit	A food	computer
Something you listen to	music	Used to play sports	kangaroo
Something you listen to	radio	A feature of a house	CD
A vehicle	train	Part of a house	arrow
A computer part	mouse	A place	mother
A thing you read	letter	A family member	sand
A hot drink	coffee	A family member	page

Appendix G

In this task, you will see 2 sets of text on the screen for each item. First, you will see a phrase. These words are a category. Then, those words will go away and you will see a single word. You need to decide if the single word fits the category or not. Press the GREEN button if the word fits the category. Press the RED button if it does not fit the category
Press the SPACEBAR to continue.

Next, you will see some examples. Remember, press the GREEN button if the word fits the category, and press the RED button if it does not fit the category. Try to press the buttons as quickly and accurately as you can.
Press the SPACEBAR to continue.

You will see 2 examples. Remember, first you will see a category, then you will see a word. Press the GREEN button if the word is something that fits in the category, and press the RED button if it does not fit.
Press the SPACEBAR to go to the first example.

A pet [1500ms - then disappear] / dog

In this example, the word 'dog' fits the category 'a pet', so you should have pressed the GREEN (Yes) button.
Press the SPACEBAR to do another example.

A place [1500ms - then disappear] / egg

In this example, 'egg' does not fit the category 'a place', so you should have pressed the RED (No) button.
Press the SPACEBAR to do 10 practice questions.

Ready? Here we go. Press the SPACEBAR to start the practice.
[Practice items]
A kind of clothing / sweater
A food / bread
A family member / bicycle
A drink / juice
Something you listen to / paper
A type of transportation / car
An electronic device / orange
Something you watch / TV
A person / desk
A food / mother

References

Aaron, P., Joshi, R., Ayotollah, M., Ellsberry, A., Henderson, J., & Lindsey, K. (1999). Decoding and sight-word naming: Are they independent components of word recognition skill? *Reading and Writing, 11*(2), 89-127.

Akamatsu, N. (1999). The effects of first language orthographic features on word recognition processing in English as a second language. *Reading and Writing, 11*(4), 381-403.

Akamatsu, N. (2003). The effects of first language orthographic features on second language reading in text. *Language learning, 53*(2), 207-231.

Akamatsu, N. (2005). Effects of second language reading proficiency and first language orthography on second language word recognition, In V.J. Cook and B. Bassetti, *Second Language Writing Systems,* (pp. 238-259), Clevedon, England: Multilingual Matters.

Akamatsu, N. (2007). The effects of training on automatization of word recognition in English as a foreign language. *Applied Psycholinguistics, 29*(2), 175-193.

Baldauf, R., & Dawson, R. (1980). The predictive validity of the Michigan test of English language proficiency for teacher trainees in Papua New Guinea. *Educational and Psychological Measurement, 40*(4), 1201-1205.

Balota, D., & Chumbley, J. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception and Performance, 10*(3), 340-357.

Baluch, B., & Besner, D. (1991). Visual word recognition: Evidence for strategic control of lexical and nonlexical routines in oral reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*(4), 644-652.

Berndt, R., Reggia, J., & Mitchum, C. (1987). Empirically derived probabilities for grapheme-to-phoneme correspondences in English. *Behavior research methods, instruments & computers, 19*(1), 1-9.

Brown, T., & Haynes, M. (1985). Literacy background and reading development in a second language. In T.H. Carr (Ed.), *The development of reading skills*, (pp. 19-34). San Francisco: Jossey-Bass.

Chen, H., & Flores, A. GB, & Cheung, SL (1995). Orthographic and phonological activation in recognizing Chinese characters. *Psychological Research, 58*, 144-153.

Chikamatsu, N. (1996). The Effects of L1 Orthography on L2 Word Recognition: A Study of American and Chinese Learners of Japanese. *Studies in Second Language Acquisition, 18*(4), 403-432.

Chikamatsu, N. (2006). Developmental Word Recognition: A Study of L1 English Readers of L2 Japanese. *The Modern Language Journal, 90*(1), 67-85.

Cho, J., & Chen, H. (1999). Orthographic and phonological activation in the semantic processing of Korean Hanja and Hangul. *Language and Cognitive Processes, 14*(5), 481-502.

Corrigan, A., Dobson, B., Kellman, E., Spaan, M., & Tyma, S. (1979). Michigan test of English language proficiency (Form P). Ann Arbor: *Michigan University Press.*

DeFrancis, J. (1989). *Visible speech: The diverse oneness of writing systems*, Honolulu: University of Hawaii Press.

Doehring, D. (1976). Acquisition of rapid reading responses. *Monographs of the society for research in child development*, 1-54.

Durgunoglu, A., Nagy, W., & Hancin-Bhatt, B. (1993). Cross-language transfer of phonological awareness. *Journal of Educational Psychology, 85*(3), 453-465.

Ehri, L. (1995). Phases of development in learning to read words by sight. *Journal of Research in Reading, 18*(2), 116-125.

Ehri, L. (1998). Grapheme-phoneme knowledge is essential for learning to read words in English. In J.L. Metsala and L. Ehri (Eds.), *Word recognition in beginning literacy*, (pp3-40). Hillsdale, NJ: Erlbaum.

Ehri, L., & Wilce, L. (1982). Recognition of spellings printed in lower and mixed case: Evidence for orthographic images. *Journal of Reading Behavior, 14*(3), 219-230.

Ellis, N. C., & Hooper, A. M. (2001). Why learning to read is easier in Welsh than in English: Orthographic transparency effects evinced with frequency-matched tests. *Applied Psycholinguistics, 22,* 571–599.

Everson, M. (1998). Word recognition among learners of Chinese as a foreign language: Investigating the relationship between naming and knowing. *Modern Language Journal*, 194-204.

Fender, M. (2003). English word recognition and word integration skills of native Arabic-and Japanese-speaking learners of English as a second language. *Applied Psycholinguistics, 24*(02), 289-315.

Fisher, D. (1975). Reading and visual search. *Memory and Cognition, 3*, 188-196.

Forster, K., & Chambers, S. (1973). Lexical Access and Naming Time. *Journal of Verbal Learning and Verbal Behavior, 12*(6), 627-635.

Forster, K., & Forster, J. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods Instruments and Computers, 35*(1), 116-124.

Freebody, P., & Byrne, B. (1988). Word-reading strategies in elementary school children: Relations to comprehension, reading time, and phonemic awareness. *Reading Research Quarterly*, 441-453.

Freed, B., Dewey, D., Segalowitz, N., & Halter, R. (2004). The language contact profile. *Studies in Second Language Acquisition, 26*(02), 349-356.

Frost, R. (1994). Prelexical and Postlexical Strategies in Reading: Evidence From a Deep and a Shallow Orthography. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(1), 116-129.

Frost, R. (1998). Toward a strong phonological theory of visual word recognition: True issues and false trails. *Psychological Bulletin, 123*, 71-99.

Frost, R., & Katz, L. (1992). *Orthography, phonology, morphology, and meaning*, Amsterdam: North-Holland.

Frost, R., Katz, L., & Bentin, S. (1987). Strategies for visual word recognition and orthographical depth: A multilingual comparison. *Strategies, 13*(1),104-115.

Fukkink, R., Hulstijn, J., & Simis, A. (2005). Does training in second-language word recognition skills affect reading comprehension? An experimental study. *Modern Language Journal*, 89(1),54-75.

Geva, E., & Siegel, L. (2000). Orthographic and cognitive factors in the concurrent development of basic reading skills in two languages. *Reading and Writing, 12*(1), 1-30.

Goswami, U., Gombert, J., & de Barrera, L. (1998). Children's orthographic representations and linguistic transparency: Nonsense word reading in English, French, and Spanish. *Applied Psycholinguistics, 19*(1), 19-52.

Gough, P. (1984). Word recognition. *Handbook of reading research, 1*, 225-253.

Green, D., & Meara, P. (1987). The effects of script on visual search. *Second Language Research, 3*(2), 102.

Hino, Y., & Lupker, S. (1998). The effects of word frequency for Japanese Kana and Kanji words in naming and lexical decision: Can the dual-route model save the lexical-selection account? *Journal of Experimental Psychology, Human Perception, and Perforamnce, 24*, 1431-1453.

Hino, Y., & Lupker, S. (2000). Effects of word frequency and spelling-to-sound regularity in naming with and without preceding lexical decision. *Journal of experimental psychology. Human perception and performance, 26*(1), 166.

Holm, A., & Dodd, B. (1996). The effect of first written language on the acquisition of English literacy. *Cognition, 59*(2), 119-147.

Holmes, V., & Carruthers, J. (1998). The relation between reading and spelling in skilled adult readers. *Journal of Memory and Language, 39*(2), 264-289.

Hoosain, R. (1991). *Psycholinguistic implications for linguistic relativity: A case study of Chinese*: L. Erlbaum Associates.

Jared, D. (2002). Spelling-sound consistency and regularity effects in word naming. *Journal of Memory and Language, 46*(4), 723-750.

Jiang, N. (1999). Testing processing explanations for the asymmetry in masked cross-language priming. *Bilingualism: Language and Cognition, 2*(01), 59-75.

Jiang, N. (2007). Selective integration of linguistic knowledge in adult second language learning. *Language Learning, 57*(1), 1-33.

Juffs, A. (2001). Psycholinguistically Oriented Second Language Research. *Annual Review of Applied Linguistics, 21*, 207-220.

Katz, L., & Feldman, L. (1983). Relation between pronunciation and recognition of printed words in deep and shallow orthographies. *Relation, 9*(1), 157-166.

Katz, L., & Frost, R. (1992). The reading process is different for different orthographies: The orthographic depth hypothesis. *Orthography, phonology, morphology, and meaning*, 67-84.

Kessler, B., & Treiman, R. (2001). Relationships between sounds and letters in English monosyllables. *Journal of Memory and Language, 44*(4), 592-617.

Kinoshita, S. (1987). Case alternation effect: Two types of word recognition? *The Quarterly Journal of Experimental Psychology Section A, 39*(4), 701-720.

Koda, K. (1988). Cognitive process in second language reading: Transfer of L1 reading skills and strategies. *Second Language Research, 4*(2), 133.

Koda, K. (1989). Effects of L1 orthographic representation on L2 phonological coding strategies. *Journal of Psycholinguistic Research, 18*(2), 201-222.

Koda, K. (1990). The Use of L1 Reading Strategies in L2 Reading: Effects of L1 Orthographic Structures in L2 Phonological Recoding Strategies. *Studies in Second Language Acquisition, 12*(4), 393-410.

Koda, K. (1999). Development of L2 intraword orthographic sensitivity and decoding skills. *Modern Language Journal*, 51-64.

Koda, K. (2005). *Insights into second language reading: A cross-linguistic approach*, Cambridge: Cambridge University Press.

Kucera, H., & Francis, W. (1967). *Computational analysis of present-day American*

English, Providence: Brown University Press.

LaBerge, D., & Samuels, S. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology, 6*(2), 293-323.

Leong, C. (1986). What does accessing a morphemic script tell us about reading and reading disorders in an alphabetic script? *Annals of Dyslexia, 36*(1), 82-102.

Lesch, M., & Pollatsek, A. (1998). Evidence for the use of assembled phonology in accessing the meaning of printed words. *Journal of experimental psychology. Learning, memory, and cognition, 24*(3), 573.

Luo, C. (1996). How is word meaning accessed in reading? Evidence from the phonologically mediated intereference effect. *Journal of experimental psychology. Learning, memory, and cognition, 22*(4), 883-895.

McCann, R., Folk, C., & Johnston, J. (1992). The role of spatial attention in visual word processing. *Journal of Experimental Psychology, 18*(4), 1015-1029.

Midgley, K., Holcomb, P., & Grainger, J. (2009). Language effects in second language learners and proficient bilinguals investigated with event-related potentials. *Journal of Neurolinguistics, 22*(3), 281-300.

Mori, Y. (1998). Effects of first language and phonological accessibility on kanji recognition. *Modern Language Journal*, 69-82.

Muljani, D., Koda, K., & Moates, D. (1998). The Development of Word Recognition in a Second Language. *Applied Psycholinguistics, 19*(1), 99-113.

Oney, B., & Durgunoglu, A. (1997). Beginning To Read in Turkish: A Phonologically Transparent Orthography. *Applied Psycholinguistics, 18*(1), 1-15.

Paap, K., Newsome, S., & Noel, R. (1984). Word shape's in poor shape for the race to the lexicon. *Journal of Experimental Psychology: Human Perception and Performance, 10*(3), 413-428.

Paap, K., & Noel, R. (1991). Dual-route models of print to sound: Still a good horse race. *Psychological Research, 53*(1), 13-24.

Peereman, R., Content, A., & Bonin, P. (1998). Is perception a two-way street? The case of feedback consistency in visual word recognition. *Journal of Memory and Language, 39*, 151-174.

Perfetti, C. (2003). The universal grammar of reading. *Scientific Studies of Reading, 7*(1), 3-24.

Perfetti, C., Liu, Y., Fiez, J., Nelson, J., Bolger, D., & Tan, L. (2007). Reading in two writing systems: Accommodation and assimilation of the brain's reading network. *Bilingualism: Language and Cognition, 10*(02), 131-146.

Perfetti, C., Liu, Y., & Tan, L. (2005). The lexical constituency model: Some implications of research on Chinese for general theories of reading. *Psychological Review, 112*(1), 43-59.

Perfetti, C., & Zhang, S. (1995). Very early phonological activation in Chinese reading. *Learning, Memory, 21*(1), 24-33.

Perfetti, C., Zhang, S., & Berent, I. (1992). Reading in English and Chinese: Evidence for a universal phonological principle. *Orthography, phonology, morphology, and meaning, 94*, 227-248.

Plaut, D., & McClelland, J. „Seidenberg, MS, & Patterson, ?.(1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review, 103*, 56-115.

Preacher, K., Rucker, D., MacCallum, R., & Nicewander, W. (2005). Use of the extreme groups approach: A critical reexamination and new recommendations. *Psychological methods, 10*(2), 178-192.

Rumsey, J., Horwitz, B., Donohue, B., Nace, K., Maisog, J., & Andreason, P. (1997). Phonological and orthographic components of word recognition. A PET-rCBF study. *Brain, 120*(5), 739-759.

Schneider, W., & Shiffrin, R. (1977). Controlled and Automatic Human Information Processing: 1. Detection, Search, and Attention. *Psychological Review, 84*(1), 1-66.

Segalowitz, N. (1986). Skilled reading in the second language.In J. Vaid (Ed.), *Language processing in bilinguals: Psycholinguistic and neuropsychological perspectives*, (pp.3-19). Hillsdale, NJ: Lawrence Erlbaum.

Segalowitz, N., & Hulstijn, J. (2005). Automaticity in bilingualism and second language learning. *Handbook of bilingualism: Psycholinguistic approaches*, 371-388.

Segalowitz, N., & Segalowitz, S. (1993). Skilled Performance, Practice, and the Differentiation of Speed-Up from Automatization Effects: Evidence from Second Language Word Recognition. *Applied Psycholinguistics, 14*(3), 369-385.

Segalowitz, S., Segalowitz, N., & Wood, A. (1998). Assessing the Development of Automaticity in Second Language Word Recognition. *Applied Psycholinguistics, 19*(1), 53-67.

Seidenberg, M. (1988). Cognitive neuropsychology and language: The state of the art. *Cognitive Neuropsychology, 5*(4), 403-426.

Seidenberg, M. (1992). Connectionism without tears. In S. Davis (Ed.), *Connectionism: Theory and practice*, (84-122). New York: Oxford University Press.

Seymour, P., Aro, M., & Erskine, J. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology, 94*(2), 143-174.

Spencer, A. (1991). *Morphological theory: An introduction to word structure in generative grammar*: Oxford: Blackwell.

Stanovich, K. (1980). Toward an interactive-compensatory model of individual differences in the development of reading fluency. *Reading research quarterly, 16*(1), 32-71.

Stanovich, K. (1991). Word recognition: Changing perspectives. *Handbook of reading research, 2*, 418-452.

Tabossi, P., & Laghi, L. (1992). Semantic priming in the pronunciation ofwords in two writing systems: Italian and English. *Memory & Cognition, 20*(3), 303-313.

Tan, L., & Perfetti, C. (1997). Visual Chinese character recognition: Does phonological information mediate access to meaning? *Journal of Memory and Language, 37*(1), 41-57.

Tan, L., & Perfetti, C. (1998). Phonological codes as early sources of constraint in Chinese word identification: A review of current discoveries and theoretical accounts. *Reading and Writing, 10*(3), 165-200.

Tan, L., & Perfetti, C. (1999). Phonological activation in visual identification of Chinese two-character words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*(2), 382-393.

Tan, L., Spinks, J., Gao, J., Liu, H., Perfetti, C., Xiong, J., et al. (2000). Brain activation in the processing of Chinese characters and words: a functional MRI study. *Human Brain Mapping, 10*(1), 16-27.

Van Orden, G. (1987). A ROWS is a ROSE: spelling, sound, and reading. *Memory & Cognition, 15*(3), 181-198.

Wade-Woolley, L. (1999). First language influences on second language word reading: All roads lead to Rome. *Language Learning, 49*(3), 447-471.

Wang, M., & Geva, E. (2003). Spelling performance of Chinese children using English as a second language: Lexical and visual–orthographic processes. *Applied Psycholinguistics, 24*(01), 1-25.

Wang, M., Koda, K., & Perfetti, C. (2003). Alphabetic and non-alphabetic L1 effects in English semantic processing: A comparison of Korean and Chinese English L2 learners. *Cognition, 87*, 129-149.

Wydell, T., Patterson, K., & Humphreys, G. (1993). Phonologically mediated access to meaning for Kanji: is a Rows still a Rose in Japanese Kanji? *Journal of experimental psychology. Learning, memory, and cognition, 19*(3), 491-514.

Yin, B., & Felley, M. (1990). *Chinese Romanization: Pronunciation and Orthography*: Beijing: Sinolingua.