CrossMark

# The Resolution of Visual Noise in Word Recognition

**Hye K. Pae · Yong-Won Lee**

**Abstract** This study examined lexical processing in English by native speakers of Korean and Chinese, compared to that of native speakers of English, using normal, alternated, and inverse fonts. Sixty four adult students participated in a lexical decision task. The findings demonstrated similarities and differences in accuracy and latency among the three L1 groups. The participants, regardless of L1, had a greater advantage in nonwords than words for the normal fonts because they were able to efficiently detect the illegal letter strings. However, word advantages were observed in the visually distorted stimuli (i.e., alternated and inverse fonts). These results were explained from the perspectives of the theory of psycholinguistic grain size, L1–L2 distance, and the mechanism of familiarity discrimination. The native speakers of Chinese were more sensitive to visual distortions than the Korean counterpart, suggesting that the linguistic template established in L1 might play a role in word processing in English.

**Keywords** Lexical processing · Alternated fonts · Inverse fonts · Psycholinguistic grain size theory

## Introduction

Reading is a process that converts the print to the equivalent sound in order to gain comprehension. Beyond this universality, each language has a minimal processing unit (i.e., grain size) that is utilized in the conversion route from a string of letters to the corresponding sounds (Ziegler and Goswami 2005). Along with the grain size (e.g., graphemes/phonemes, onsets, rimes, syllables; Ziegler and Goswami 2005), the phonotactic rules that govern the

H. K. Pae (✉)
School of Education, University of Cincinnati, Cincinnati, OH, USA
e-mail: hye.pae@uc.edu

Y.-W. Lee
Department of English Language and Literature, Seoul National University, Seoul, South Korea
e-mail: ylee01@snu.ac.kr

spelling pattern also play a part in reading processes. The examination of word recognition provides not only an understanding of the developmental aspect of proficient reading, but also an explanation of the role of grain sizes and phonotactic rules in word reading, because it serves as a window into the intricacy of letter-string constituents involved in reading. Given the abundant evidence of English word recognition by native speakers, it would be useful to see how speakers of English as a foreign language (FL) and a second language (L2[1]) who have different first language (L1) backgrounds process English words and nonwords in various visual font shapes. This study explores the pattern of lexical recognition and processing speed in Korean- and Chinese-speaking college students who learned English as FL in their native countries, in comparison to that of native English speakers, utilizing a lexical decision task. Since both L2 groups have distinct L1s (i.e., alphasyllabary vs. syllabic logography), this study will make vertical and horizontal advances in the field of L1 influences on L2 English lexical processing.

Orthographic Depth and Word Processing

Depending on the representation of symbols, the writing system is viewed as glottography or logography. Glottographic languages use symbols that represent sounds in the form of either phonemes in alphabetic languages (e.g., English, Korean) or syllables in syllabary (e.g., Japanese Kana), whereas logographic writing systems use symbols that represent "ideas" (a.k.a., ideography; e.g., Chinese, Japanese Kanji). The writing system is closely related to orthography, which is characterized by the depth of orthography with varying degrees across languages (Frost et al. 1987). It is largely due to the notion that learning to read depends on how a writing system encodes oral language (Perfetti and Liu 2005). The degree of transparency in the relationship between graphemes (as in alphabetic languages) or syllables (as in syllabic languages) and the minimal sound unit (phonemes in alphabetic languages and syllables in syllabic languages) varies across languages. In deep orthographies in alphabetic languages, such as English, a grapheme at times does not map onto a single phoneme, as one grapheme has more than one sound (i.e., one-to-many correspondence); consequently, the sound of vowels and consonants is irregular and inconsistent (e.g., hear-bear, nation-national, lead-lead, church-monarch, etc.). In shallow orthographies, such as Spanish, Italian, and Korean, in contrast, a letter has a regular and consistent alignment with its sound, as one grapheme has one sound (i.e., one-to-one correspondence).

The orthographic depth hypothesis has been a conceptual framework that explains differences in word processing across different linguistic systems (Frost 1998; Goswami et al. 2003), because the effect of orthographic depth on reading is found to be variable according to the linguistic system under consideration. This variability stems from the degree of consistency to which lexical information is involved in pronunciation (Frost et al. 1987). Although the consistency in letter-sound mappings has been highlighted in the field of word recognition, there is some controversy surrounding the orthographic depth hypothesis. Specifically, Seidenberg (2012) asserts that language and writing systems need to be understood in terms of multiple constraints related to linguistic activities and biological bases as well as the interface of orthographic depth and morphological complexity. Moreover, the orthographic depth hypothesis mainly explains letter-sound mappings in alphabetic languages, as multiple graphemes are involved in the formation of a syllable in those languages. However, in syllabic languages, such as Chinese and Japanese, a character represents a syllable.

---

[1] Although learning L2 is a little different from learning FL in a strict sense because of the different learning context, L2 and FL are used interchangeably in this paper.

Psycholinguistic Grain Size Theory: Different Processing Units in English, Korean, and Chinese

Since it is sensitive to the degree to which a written word deviates from the regularity of one-to-one letter-sound correspondence, the orthographic depth hypothesis is closely related to a minimal grain size as the language-specific unit of processing in a language. According to the psycholinguistic grain size theory (Ziegler and Goswami 2005), lexical organizations and processing strategies in different orthographies are influenced by the granularity defined by the writing system. In alphabetic languages, the minimal processing unit is a grapheme/phoneme. In syllabic languages, however, the minimal grain size is a syllable. The different processing unit in English, Korean, or Chinese is characterized by the grain size of the given language. The psycholinguistic grain size theory posits that English (deep orthography) speakers use bigger and more grain sizes than Korean (shallow orthography) speakers, because a reliance on a small grain size in English leads to incorrect pronunciation, resulting from the irregularity between graphemes and phonemes. Due to the syllabic nature of the Chinese language, the grain size of Chinese is basically a syllable.[2]

The hierarchical granularity in English refers to top-down segmentation from whole words, syllables, onsets/rimes to individual letters (Ziegler and Goswami 2005). It seems that sub-syllabic units are language universal at least in alphabetic languages, but processing units are language specific, suggesting that the grapheme–phoneme interface affects the processing unit of a particular language. Previous research supports the claim that there are different patterns of segmenting intraword units across languages (Wang and Cheng 2008; Yoon et al. 2002; Ziegler and Goswami 2005). In a segmentation task, English speakers tend to divide words into onset-rime units (e.g., the C-VC unit) rather than body-coda units (e.g., the CV-C unit). For example, the word "pat" is phonetically divided by English-speaking children into an onset "p" and a rime "at" (i.e., /p/+/æt/) rather than a body "pa" and a coda "t" (i.e., /pæ/+/t/). However, this is not universal across languages. Yoon et al. (2002) show that Korean speakers demonstrate preference for the body-coda unit as an orthographic chunk over the onset-rime, when they are asked to segment the speech sound in words. For example, Korean children are likely to segment the word "pat" into a body "pa" and a coda "t" (i.e., /pæ/+/t/). This sharp difference allows for a logical inference that a unique processing mechanism is involved in visual word recognition in a particular language.

The effects of syllabic and phonemic awareness skills on word identification are dependent on the linguistic system. For instance, Cho and McBride-Chang (2005) also showed that, although phonemic awareness skills were a significant predictor of English word reading by Korean children, syllabic awareness was a stronger predictor than phonemic awareness in Korean word reading. It seems that Korean children do not develop phoneme sensitivity as much as English monolinguals do due to the salient syllabic structure of Korean and shallow orthographic features. In addition, Simpson and Kang (2004) found, in a naming test of Korean, that the writing convention of Korean in syllable blocks affects syllable processing beyond lexical and subsyllabic properties. They found that, when syllable frequency was controlled for, free syllables and bound syllables had similar naming speed, but pseudosyllables were named slower than bound syllables whose frequency effects were substantial. Given that Korean is a shallow orthography (i.e., regular letter-sound correspondence), the underlying assumption was no frequency effect. This finding suggests that syllables serve as a functional unit in alphabetic Korean words. These results reflect the difference of linguistic

---

[2] We refer to the canonical aspect of the Chinese language which is syllabic. Hence, the artificial *Pinyin* is beyond our discussion here.

systems between English and Korean. Although English and Korean share the alphabetic principle, Korean is different from English in the degree of syllabic dominance. Korean is an alphasyllabic language that adheres to the alphabetic principle and, at the same time, has a dominant syllabic feature (see Pae 2011, for more information). Like English, a phoneme maps onto a grapheme and multiple graphemes group together to form a syllable (i.e., the alphabetic principle). Unlike English, however, a string of graphemes is not written in a horizontally linear form but in a distinct syllabic block (e.g., *Korean* is written as 한글 rather than ㅎㅏㄴㄱㅡㄹ). In the visual form, Korean is closer to Chinese than English because of the grapheme package in block (e.g., 韓國語 meaning "the Korean language" in Chinese character). Chinese is known as a logographic, syllabic language in which a letter corresponds to a syllable. For example, simple strokes, as seen in "一," "二" meaning "one", "two", respectively, are in conformity to syllables.

Research shows that intrasyllabic unit processing in Chinese is different from that in English. For instance, Wang and Cheng (2008) probed subsyllabic-unit processing in young Chinese natives and Chinese–English bilingual children, using a sound similarity judgment task and a sound matching task. They found that Chinese native children showed a shared body-unit primacy, whereas bilingual children preferred the body over the rime in Chinese, but preferred the rime over the body in English. They have concluded that different intraword units are involved in processing different linguistic systems.

Research evidence has also demonstrated that syllable awareness is the strongest predictor of word recognition and that phonemic onset awareness plays a weak role in Chinese character recognition in Chinese young children (McBride-Chang et al. 2004). These language-specific findings are attributable to the language system in which Chinese does not require onset-rime awareness, because each character maps onto a spoken syllable (McBride-Chang et al. 2004). In another study of Chinese young children's word recognition in Chinese and English, McBride-Chang et al. (2008) found that syllable awareness equally accounted for the significant variance in both Chinese and English word recognition. Syllable onset awareness was uniquely related to English reading, while tone detection served as a crucial predictor of Chinese reading. The authors interpreted these results as support for the universality of L1 phonological transfer to L2 reading as well as different psycholinguistic grain sizes involved in word reading.

Monosyllabic- and multisyllabic-word recognition has also been investigated. Chen and Vaid (2007), for instance, attempted to determine whether native English readers segmented polysyllabic words based on orthographic and morphological criteria or phonological information. Based on the results showing that a primacy for the basic orthographic syllable structure was consistently found in low frequency words but not in high frequency ones, they concluded that word frequency modulated the segmentation strategy of polysyllabic English words. Concerning the locus of frequency effects, Lee (1999) has argued that word-length and word-frequency effects have a shared locus in word processing, as word frequency and word-length effects are generated at a prelexical or lexical-access stage

L1 Influence on L2 Word Recognition

Cross-language transfer has been extensively discussed in the recent decades. The interconnection between L1 and L2 has been emphasized from phonological awareness to reading comprehension. A number of studies showed that phonological sensitivity was a necessary condition, if not sufficient, for efficient word reading in L1 and L2 (Durgunoglu and Oney 1999; Gottardo et al. 2001; McBride-Chang et al. 2004; Pae et al. 2010).

L1 influence on L2 word decoding has been investigated from different angles, such as phonological and orthographic role in word identification, vowel sensitivity in L1 and L2, and effects of L1 orthographic features on L2 reading. Learners of English as L2 with different L1 backgrounds (i.e., alphabetic vs. nonalphabetic) tend to demonstrate differences in the reliance of phonological and orthographic information on visual word recognition. Wang et al. (2003) tested two groups of Chinese (nonalphabetic) and Korean (alphabetic) adult learners of English using a categorization judgment task in English. The participant was asked to determine whether a noun (e.g., *rose* or *brain*) presented on the computer screen was a part of a certain category (e.g., *flower* or *body part*), in which a homophonic stimulus was used for the correct exemplar (e.g., *rows* for *rose*, *brane* for *brain*). Korean students' performance was influenced more by phonological information than orthographic properties by making more errors on homophonic targets (e.g., *rows*, *brane*) than on control stimuli. In contrast, Chinese students showed a heavier reliance on orthographic information than the phonological feature, in which they showed a significant difference in error rates between orthographically similar and less similar homophonic targets. These results indicated that orthographic information played a significant role in their processing of English words. Their findings suggested that different L1 systems and their corresponding L1 processing mechanisms were the source of differences in L2 lexical processing and word identification. In other words, L2 learners with an alphabetic L1 tend to rely on phonological information more than L2 learners with a nonalphabetic L1 whose processing is more sensitive to orthographic properties. The results of McBride-Chang et al.'s (2008) study also suggest more phonological involvement in reading an alphabetic language than a logographic language. These results converge on a possibility of different lexical processing between alphabetic and nonalphabetic languages.

Moreover, Akamatsu (1999, 2003) investigated the effect of the orthographic feature of three different L1s on L2 English word recognition, and found different levels of L1 effects on L2 English reading. Akamatsu (1999, 2003) compared the reading performance of Iranians, Chinese, and Japanese speakers to that of native English speakers. On the basis of the results that Iranian students who spoke alphabetic L1 (Persian) were less affected by the CaSe AlTeRnAtIoN (i.e., visual noise) than Chinese and Japanese counterparts who spoke syllabic languages, Akamatsu (1999, 2003) concluded that L1 orthographic features influenced L2 English reading. In a similar study, a greater mixed-case disadvantage for nonwords than words was found for a longer duration of stimulus presentation (Allen et al. 1995). When targets were presented for a shorter duration followed by masks, the result was reversed; that is, word targets showed a larger mixed-case disadvantage. Word frequency effects were found in mixed fonts.

Research results indicated that case alternation effects were stronger in real words than nonwords in a lexical decision task (Besner 1983; Lavidor et al. 2002). Besner (1983) and Lavidor et al. (2002) interpreted this phenomenon as the 'familiarity discrimination mechanism', viewing the figural pattern of words as an estimate of the stimulus visual familiarity.

The Present Study

In order to investigate how visual noise is resolved in word recognition by nonnative speakers of English, three font shapes were utilized in this study: normal fonts (e.g., English), alternated fonts (e.g., eNgLiSh), and upside-down fonts (e.g., ɥsıꞁɓuƎ). The alternated font was used as intraword noise (i.e., visual information within the lexicon is distorted) in this study, because significant effects of lexical shapes have been reported in previous research studies (Akamatsu 1999, 2003; Allen et al. 1995; Besner and McCann 1987; Reingold et al. 2010). The inverse font was utilized as holistic noise (i.e., the intraword visual information is preserved, but

the whole lexical part is inverted) in this study, because (1) there was no published article that utilized upside-down fonts in the literature of word recognition by Korean and Chinese learners of English and (2) English words carry cuing information differently in the upper and bottom parts of print.

The English alphabets are visually more prominent in the upper part than the bottom part. As seen below, about twice as many alphabet letters ascend above the median, and most of the ascending letters are consonants. If English words were upside down, the properties of phonotactic rules and cuing information would be lost.

abcdefghijklmnopqrstuvwxyz

As a result, in a visual presentation, most English readers read bottom-part-absent sentences faster than top-part-absent ones (e.g., I like apples vs. I like apples) because the upper part of prints holds more meaningful information than the bottom part (Weaver 1988). Moreover, English has a unique phonotactic rule that characterizes the plausible sequence of letters in words. The phonotactic principle constrains permissible consonant clusters and syllable structures from a finite set of letter strings. For example, the letter string *the* is legal, while *tqe* is illegal. In three-letter words, *h* follows *t* about ten times more often than any other onset letter, and *h* precedes *e* about twenty times more often than any other letter (Adams 1990). Upon gaining automaticity of word recognition, skillful readers make use of the phonotactic rule quickly, accurately, and effortlessly by efficiently eliminating illegal strings when they process the text. Onsets of words carry phonotactic rules to a great extent. Endings are more predictable than the onset and middle positions because they bear grammatical information (e.g., *-ness*: nouns; *-ful*: adjectives). Middles typically contain vowels (e.g., c*a*t, p*a*t). If English words were inverted, the left-to-right phonotactic rules are disoriented and typical cuing information would be reduced.

Given these differences found in cross-language studies, research into Korean and Chinese speakers' word recognition of English, in contrast to that of English speakers, provides valuable information, due to the unique properties that the three languages entail. In order to examine how the L1 system affects L2 word recognition of nonnative speakers of English, Korean- and Chinese-speaking college students were recruited. The two groups were selected because they have different L1 writing systems (a glottographic alphasyllabary vs. a logographic, syllabic language) and because the adults have already gained automaticity in English word decoding. Three research questions were examined in this study.

1. How is the performance of Korean and Chinese nonnative speakers of English (i.e., accuracy and latency) on a lexical decision test different from that of the native speakers?

   • Working Hypothesis: Given that Korean shares alphabetic characteristics with English, we hypothesized that the Korean participants' performance would be more likely to be similar to that of the native English speakers than Chinese participants, although there would be similarities and differences in the performance pattern of the two groups.

2. How do the three L1 groups perform on the different visual shapes (i.e., normal, alternated, and inverse fonts) with respect to lexical properties, such as base-word frequency, lexical variation, and lexical length?

   • Working Hypothesis: Given the L1-L2 distance, we hypothesized that the Chinese participants would show greater noise effects than the Korean counterpart.

3. To what extent do lexical frequency and lexical length play a role in word recognition with the three visual shapes (i.e., normal, alternated, and inverse fonts)?

- Working Hypothesis: Given the different grain sizes in Korean and Chinese, we hypothesized that the Chinese participants would be more susceptible to visual disruptions, especially to the inverse fonts.

## Methods

Participants

Three language groups of 64 college students participated in this study. Although their L1s were different from each other, Korean-speaking and Chinese-speaking participants were fluent readers of English, and they had normal vision or corrected vision within the normal range. The first group was composed of 26 Korean-speaking English majors (24 undergraduate and 2 graduate students) who were residing and learning English as FL at one of the National Universities in South Korea. The second group consisted of 20 Chinese adult students who first learned English as FL in China and were continuing to learn English as L2 while learning contents of majors in the U.S. The third group included 18 English-speaking college students. The last two groups were composed of undergraduate and graduate students at a large university in the Midwestern area of the U.S. The mean age of the Korean participants were 21.65 (SD 2.65, range 19–30), while that of the Chinese students was 25.25 (SD 2.65, range 21–31). The native speakers' mean age was 28.12 (SD 5.83, range 21–38). Female students comprised of 85 % (22 examinees) of the Korean participants. The Chinese students included 65 % females and 35 % males. The native-speaker group included 12 females and 6 males. For the Korean participants, the mean year of residence in English-speaking countries was 1.94 (SD 4.04, range 0–14). For the Chinese counterpart, the mean year of residence in the U.S. was 1.49 (SD 1.56, range 0–5).

The Korean and Chinese participants' self-rating of English skills[3] was also obtained. The Korean and Chinese groups' self-rating on a scale of 1–10 (10 near-native) was remarkably similar. The means of the Korean participants' ratings were as follows (standard deviations in parentheses): speaking = 6.50 (2.16), listening = 7.00 (1.81), reading = 7.27 (1.51), and writing = 6.42 (1.58). The means of the Chinese ELs' self-ratings were as follows: speaking = 6.55 (1.32), listening = 7.00 (1.62), reading = 7.90 (1.25), and writing = 6.80 (1.51). The reading skills were rated as the highest among the four English skills for both groups, while speaking and writing were relatively lower than reading.

Procedure

A lexical decision task was utilized to obtain the three groups' accuracy and latency scores on word and nonword judgment. The participant was individually asked to make decision as

---

[3] The level of English proficiency was also gauged using the participants' proficiency test scores. The Korean participants were high-performing English majors. Some reported TOEFL scores, while others had comparable English proficiency test scores, such as the Test of English for International Communication (TOEIC) and the Test of English Proficiency developed by professionals at Seoul National University in Korea. The mean score of the TOEFL iBT test was 110.61, which was approximately equivalent to 637 on the paper-based test and the 92nd percentile. Students who had the other test scores were at the similar level to those of TOEFL, as indicated by the 94th percentile on TOEIC. The mean score of TOEFL iBT of Chinese participants was 94.29, whose equivalency of the paper-based test was 587. It should be noted that, although there was a difference in the English proficiency test scores between the Korean and Chinese groups, the difference would not prevent us from comparing the two groups in this study because the skill required to perform the given task was at the level of word reading.

quickly and accurately as possible on whether the letter string on the computer screen was a real word or a nonword by pressing a designated key on the keyboard. After granting written consent, each participant saw a series of stimuli on the computer screen in a quiet room. The stimulus was presented visually one at a time and was randomized upon presentation. The lexical decision test included 6 practice items and 60 target stimuli, and each practice or target item was preceded by a fixation point (+) presented for 500 milliseconds (ms). Each target stimulus was shown on the screen until the participant responded or it was timed out at 4,000 ms. Upon completing the task, the participants filled out a questionnaire including demographic information and self-assessed English proficiency. The accuracy and processing speed data generated by the computer were transferred to Excel and SPSS spreadsheets for analyses.

In addition, word frequency for each base word was obtained from the Corpus of Contemporary American English database (COCA; Davies 2013). COCA is a freely available online database including 425 million words from various genres, such as spoken, fiction, popular magazines, newspapers, and academic journals. As a measure of stimulus length, the numbers of letters and syllables were counted from the base word, and each variable was entered into the regression equation separately, when necessary.

### Instrument

A lexical decision task[4] including 60 stimuli was constructed for this study, using DMDX.[5] The base words for the stimuli were drawn from the Word Identification subtest of the *Woodcock Reading Mastery Test-NU* (WRMT; Woodcock 1998a). The WRMT is a well-known, U.S. norm-referenced, widely used test in reading research and schools. According to the technical manual (Woodcock 1998b), validity and reliability evidence was excellent. The split-half reliability coefficient for the college normative sample was .94 and that for adults was .97 (Woodcock 1998b). The WRMT was used in this study, because (1) it provided solid psychometric information for college students and (2) a wide range of syllables covered in the items was suitable to accomplish the objective of this study. The grade equivalency of the base words used in this study was 6th grade 7 months (Woodcock 1998b, p. 143).

The 60 base words[6] were systematically classified to construct three different visual shapes and two different lexical variations. The three different visual shapes included 20 normal fonts (e.g., 'fast'), 20 alternated fonts (e.g., 'hElP'), and 20 upside-down fonts (e.g., ɯıʌʌs), and the two different lexical forms consisted of 30 words and 30 nonwords. Nonwords were made by altering one letter of each real word (e.g., 'little' became 'luttle'). Nonwords were constructed systematically by replacing one vowel in each real word with another vowel (e.g., 'stove' became 'stuve'). The decision of altering only vowels was made for two reasons: (1) To preserve the phonotactic rule for the consonant string in the stimuli and (2) to reduce the degree of freedom in the combination of letter strings. English vowels not only make the language a deep orthography (e.g., note the vowel sound in the pair of *save-have, though*-

---

[4] Since the standard lexical decision task asks participants to read the string of letters in front of them and decide whether it is a word or not, and then to make their response by pressing a designated key, criticism has arisen. The criticism concerns more errors and longer response times due to the fact that they have to remember which key to press. An alternative way has been proposed with the go/no-go task (Perea et al. 2002). However, no significant difference has been found between the traditional way and the alternative. Since the purpose of this study was to compare the performance pattern among the three groups (i.e., the test condition was equally the same across the groups), the core of the criticism was considered insignificant.

[5] The DMDX is a freely available online: http://www.u.arizona.edu/~kforster/dmdx/download.htm. Credit should be given to Kenneth Foster and his colleagues at the University of Arizona.

[6] A DMDX syntax for the stimulus set is available from the first author.

*tou*gh, *lead-lead*, and *bear-hear*), but also appear more frequently in the text than consonants. The six basic vowels (i.e., *a, e, i, o, u, y*) take up approximately 39 % of the space of written English texts (Adams 1990). The accuracy rate and processing speed of the given stimuli were recorded and used for analyses.

## Results

Preliminary data screening in terms of missing data, score ranges, and scatterplots revealed that neither noticeable outliers nor systematic missing data were observed. Therefore, data trimming was not necessary. Before analyses were performed, data transformation took place. Percentages of correct are ordinal-level measures and fail to express scores on a linear scale. They tend to cluster scores around the center scores and do not adequately contrast the scores of the high and low ends on the scoring continuum (Bond and Fox 2007). Making a few extra-marks around the midpoint does not reflect the same score increase at the top end or at the bottom end, and a difference of ten score points at the high and low ends on the scoring continuum yields different meanings.[7] For this reason, a mathematical transformation procedure into success-to-failure ratios or odds was employed to produce linear, useful interval measures, as suggested by Bond and Fox (2007). Except the values in Table 1, the natural log values were used in analyses.

The Performance of Korean and Chinese Participants in Comparison to That of Native Speakers (Research Question 1)

Before the main analysis was run, a *t* test of the Korean and Chinese participants' performance was performed using the normal fonts as baseline data in order to ensure the comparability of the two groups. The results showed no significant difference in the accuracy score [$t(38) = .461, p > .05$], meaning that the two groups did perform similarly on word reading. This result is further explained in the discussion section.

Table 1 displays the means and standard deviations of accuracy in percentage and latency in milliseconds by language group and font type. The performance of the Korean participants was more akin to that of the native speakers than the Chinese counterpart across all modalities. The intraword noise effect (i.e., the difference in the performance between the normal fonts and alternated fonts) was the greatest in the Chinese participants, followed by the Korean counterpart. The holistic noise effect (i.e., the difference in the performance between normal fonts and inverse fonts) was in the same pattern. This finding was consistent across the accuracy and latency scores.

A one-way analysis of variance (ANOVA) was performed to examine differences in performance among the three font shapes. Because of the limitation of percentage use, as indicated earlier, mathematically transformed values were used for these analyses. The Korean participants did not show a significant difference in accuracy ($p > .05$), while they showed a significant difference in latency [$F(2, 17) = 45.26, p < .001$]. The Chinese students produced significant differences in both accuracy and latency [$F(2, 17) = 5.70, p < .01$ for accuracy; $F(2, 17) = 34.51$ for latency, $p < .001$]. The native speakers demonstrated a similar pattern to that of the Koreans with no significant difference in accuracy ($p > .05$) but a significant difference in latency [$F(2, 17) = 36.45, p < .001$].

---

[7] For example, the distance (difference = 1.9) between the scores of 88 % correct (1.99 natural log) and 98 % correct (3.89 natural log) becomes 4.75 (1.9/0.4) times greater than the distance (difference = 0.4) between the scores of 45 % correct (−0.20 natural log) and 55 % correct (0.20 natural log).

**Table 1** Descriptive statistics

| | Natives | | Koreans | | Chinese | |
|---|---|---|---|---|---|---|
| | Accuracy (%) | Latency (ms) | Accuracy (%) | Latency (ms) | Accuracy (%) | Latency (ms) |
| Alternated | 90.28 (14.52) | 1,378.08 (291.01) | 83.65 (14.11) | 1,376.47 (274.01) | 76.25 (18.27) | 2,037.55 (449.20) |
| Upside-down | 86.67 (14.58) | 1,962.93 (503.40) | 78.27 (18.22) | 1,960.88 (462.04) | 64.75 (17.36) | 2,603.25 (345.17) |
| Normal | 95.55 (7.34) | 1,090.09 (179.10) | 89.42 (19.37) | 1,004.02 (200.82) | 86.75 (17.27) | 1483.30 (389.72) |
| Intraword noise effect | 5.27 | 287.99 | 5.77 | 332.45 | 10.5 | 554.25 |
| Holistic noise effect | 8.88 | 872.84 | 11.15 | 956.86 | 22 | 1119.95 |

*SD* standard deviation; SDs in parentheses

Intraword Noise Effect = difference between the normal font and the alternated font

Holistic Noise Effect = the difference between the normal font and the upside-down font

*Subject Analysis*

The sample size of the three L1 groups was different. Therefore, the equality of variances for the variables used for the three L1 groups was checked. Levene's test of equality of error variances was not significant ($p > .05$), indicating that the assumption of variance homogeneity was tenable. The statistical significance of between-subject and within-subject differences was assessed using a $3 \times 3$ ANOVA across subjects ($F_1$) and across font types ($F_2$). The three L1 groups were used as a between-subject variable and the three different fonts as a within-subject factor. For accuracy, the main effects of the L1 groups and the font types were significant: $F_1(2, 61) = 11.79$, $p < .01$, $\eta_p^2 = .328$ for the L1 groups; $F_2(2, 61) = 47.79$, $p < .001$, $\eta_p^2 = .439$ for the font types. There was a significant interaction effect between the L1 variable and the font-type variable, indicating that the participants' L1 affected their performance on the three font types: $F(4, 61) = 3.33$, $p < .05$, $\eta_p^2 = .099$. Tukey's HSD post-hoc test revealed that there were significant differences in accuracy between the native speakers and the Korean natives ($p < .05$), between the native speakers and the Chinese participants ($p < .001$), and between the Korean natives and the Chinese participants ($p < .01$). With regard to the latency data, there were significant main effects for the L1 factor and the font-type factor: $F_1(2, 61) = 17.26$, $p < .001$, $\eta_p^2 = .361$; $F_2(2, 61) = 152.88$, $p < .001$, $\eta_p^2 = .715$. A significant interaction was also found, indicating that the speed that the participants processed the three visual shapes was affected by their L1s: $F(4, 61) = 4.176$, $p < .05$, $\eta_p^2 = .120$. A Tukey HSD post-hoc test showed that significant differences were found between the native speakers and the Chinese participants ($p < .001$) and between the native Koreans and the Chinese participants ($p < .001$). There was no significant difference between the native speakers and Korean counterparts ($p > .05$).

*Item Analysis*

A $3 \times 3$ ANOVA was performed for the item data with three L1 groups as a between-subject factor and the three font types as a within-subject variable. The Levene test for equality of error variances revealed that the assumption of variance homogeneity was not violated ($p > .05$). Concerning accuracy, there was a main effect for the L1 groups ($F_1(2, 171) = 8.45$, $p < .001$, $\eta_p^2 = .090$). A main effect was also found for the three font types ($F_2(2, 171) = 5.99$, $p < .01$, $\eta_p^2 = .065$). Tukey's HSD post-hoc test revealed that there was a significant difference between the native speakers and the Chinese participants only. There was no interaction between the L1s and font types, indicating that the participants' performance on the different fonts was independent from their L1s. With regard to the latency data, main effects were found in the L1 groups and the visual shapes: $F_1(2, 171) = 50.34$, $p < .001$, $\eta_p^2 = .371$; $F_2(2, 171) = 115.11$, $p < .000$, $\eta_p^2 = .574$. Tukey's HSD post-hoc test showed that significant differences were found between the native speakers and the Chinese participants ($p < .001$) and between the Chinese and Koreans ($p < .001$). When the two L2 groups' (i.e., Korean and Chinese participants) performance was directly compared to that of the native control group, as indicated by the post-hoc results, the difference between Chinese and natives was consistently greater than that between Koreans and natives in both accuracy and latency and across the three font types.

Pearson's correlation coefficients were obtained for the variables under examination. Table 2 shows bivariate correlations among the three groups' performance by font type. The natural log values were also used for the word frequency of occurrences, the number of letters, and the number of syllables for the sake of consistency. Word frequency was

**Table 2** Bivariate correlation matrix

|  | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| *Normal fonts* | | | | | | | | | |
| 1. Word frequency | 1 | | | | | | | | |
| 2. # Letters | −.58** | 1 | | | | | | | |
| 3. # Syllables | −.52* | .87** | 1 | | | | | | |
| 4. Koreans' accuracy | .30 | −.16 | −.11 | 1 | | | | | |
| 5. Koreans' latency | −.67** | .67* | .71** | −.22 | 1 | | | | |
| 6. Chinese' accuracy | .50* | −.31 | −.19 | .89** | −.37 | 1 | | | |
| 7. Chinese' latency | −.70** | .64** | .62** | −.25 | .88** | −.43 | 1 | | |
| 8. Natives' accuracy | .39 | −.23 | −.06 | .70** | −.24 | .92** | −.33 | 1 | |
| 9. Natives' latency | −.77* | .65** | .64** | −.48* | .80** | −.69** | .83** | −.62** | 1 |
| *Alternated fonts* | | | | | | | | | |
| 1. Word frequency | 1 | | | | | | | | |
| 2. # Letters | −.71** | 1 | | | | | | | |
| 3. # Syllables | −.67** | .91** | 1 | | | | | | |
| 4. Koreans' accuracy | .28 | −.05 | .12 | 1 | | | | | |
| 5. Koreans' latency | −.62** | .54* | .43 | −.45 | 1 | | | | |
| 6. Chinese' accuracy | .38 | −.17 | −.16 | .47 | −.27 | 1 | | | |
| 7. Chinese' latency | −.66** | .70* | .65** | −.17 | .81** | −.11 | 1 | | |
| 8. Natives' accuracy | .55* | .37* | −.28* | .36 | −.56* | .29 | −.34 | 1 | |
| 9. Natives' latency | −.65** | .71** | .65** | −.13 | .73** | −.36 | .67** | −.64** | 1 |
| *Upside-down fonts* | | | | | | | | | |
| 1. Word frequency | 1 | | | | | | | | |
| 2. # Letters | −.61** | 1 | | | | | | | |
| 3. # Syllables | −.60** | .86** | 1 | | | | | | |
| 4. Koreans' accuracy | .44 | −.46* | −.44 | 1 | | | | | |
| 5. Koreans' latency | −.45* | .70** | .63** | −.38 | 1 | | | | |
| 6. Chinese' accuracy | .42 | −.59** | −.53** | .53* | −.79** | 1 | | | |
| 7. Chinese' latency | −.39 | .73** | .66** | −.52* | −.72** | −.61** | 1 | | |
| 8. Natives' accuracy | .71** | −.62** | −.67** | .56* | −.63** | .42 | −.47* | 1 | |
| 9. Natives' latency | −.63** | .75** | .69** | −.54* | .89** | −.80** | .68 | −.70* | 1 |

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

negatively correlated with the numbers of letters and syllables as well as the latency scores, which was not surprising given the Zipf's law (i.e., the most frequent words are typically short in length; cited in Pae et al. 2012) and the direction of latency (i.e., the shorter, the better). Word frequency was consistently related to the latencies of the three font types except the Chinese students' latency score of the upside-down shapes. Word frequency was less consistently related to accuracy than to latency across the three groups by font type. The Chinese students' accuracy score was significantly correlated to word frequency only in the normal fonts ($r = .50$, $p < .05$). The native speakers' accuracy was significantly associated with word frequency in the alternated and upside-down fonts ($rs = .55$, .71, $p < .01$, respectively).

By and large, the strength of the relationship between the number of letters and the three groups' performance was greater than that of the number of syllables. Both Korean and
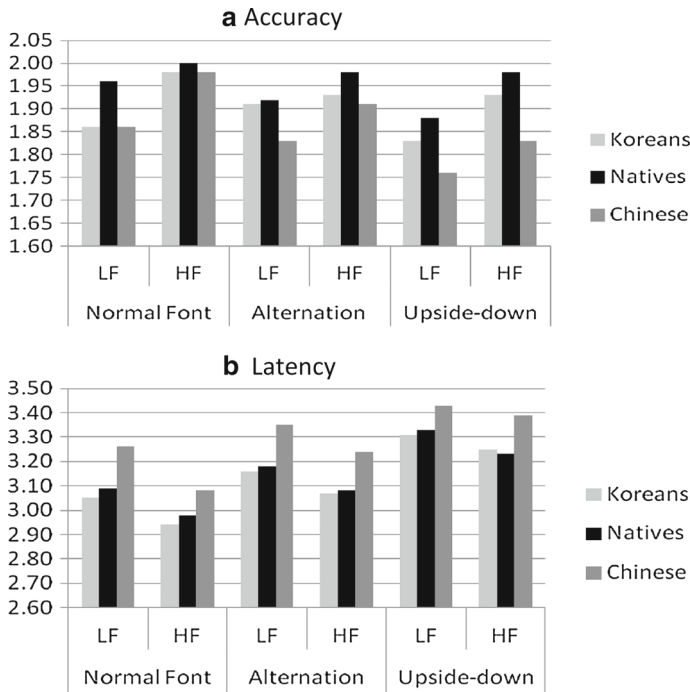
**Fig. 1** Accuracy and latency by word frequency. *LF* low frequency, *HF* high frequency

Chinese participants showed no significant correlations between accuracy and latency in the normal fonts and alternated fonts, while the native speakers showed a significant correlation between accuracy and latency in the three visual shapes ($rs = -.62$, $-.64$, and $-.70$, respectively). The Korean participants showed no significant correlation between accuracy and latency in the inverse fonts, either, whereas the Chinese counterpart showed a significant correlation in the upside-down fonts ($r = -.61$, $p < .01$).[8]

In short, the first hypothesis was supported. The Korean participants' performance was more similar to that of the native speakers than the Chinese participants.

Differences in Lexical Properties among the Three L1 Groups (Research Question 2)

In order to address Research Question 2, the three L1 groups' performance was first compared by word frequency. The influence of word frequency and the three font types on the accuracy of word recognition was examined. As seen in Fig. 1, larger disadvantages for lower frequency stimuli than higher frequency ones were found across the three groups and the three font types. The Korean and Chinese participants showed similar responses to the normal fonts ($p > .05$), but they demonstrated a significantly different pattern in the accuracy of the inverted fonts [$t(19) = 2.80$, $p < .05$]. It should be noted that the approach to the comparison under the second research question is a "minimalist's attempt" because of statistical power resulted

---

[8] Correlation coefficients between the TOEFL score and word recognition skills were obtained for the two groups. Nonsignificant correlations were found in the Korean-speaking ($r = .03$, $p > .05$) and Chinese-speaking participants ($r = .22$, $p > .05$). Since the nature of the test items used in this study was greatly different from that of TOEFL, this finding was expected.

from the small number of items of words or nonwords within each font type. In order to accommodate, effect sizes (Cohen's *d*) were obtained for the mean difference between the two L2 groups with respect to word frequency by font type, because effect size indices are independent of sample size. The effect size values showed how many standard deviations' difference there was between the means and the two groups' performance. Medium effect sizes were found in accuracy data for the low and high frequency stimuli in the inverse fonts ($d = .42$, effect size $r = .21$ for the low frequency stimuli; $d = .61$, effect size $r = .29$ for the high frequency stimuli). This means that the score of the average person in the Korean group is 0.42 standard deviations in the low frequency stimuli and 0.61 standard deviations in the high frequency stimuli above the average person in the Chinese group. However, low effect sizes for the rest of the stimulus bands were found ($d = .22$, effect size $r = .11$ for the low frequency stimuli in the alternated font; $d = .09$, effect size $r = .04$ for the high frequency stimuli in the alternated font; $d = .04$, effect size $r = .02$ for the low frequency stimuli in the normal font; $d = .03$, effect size $r = .02$ for the high frequency stimuli in the normal font). This indicates that the size of the difference in the normal font or alternated font is not so great as in the inverse font. This finding is consistent with the hypothesis set up for the second research question that the Chinese group would show greater noise effects than the Korean counterpart.

The processing speed was also compared by word frequency and font. The processing speed was in the opposite direction to accuracy, because the faster, the better. Figure 2b displays latencies taken to process the stimulus targets on the computer screen, showing that latencies are influenced by the three font types for the three L1 groups. Interestingly, the Korean participants responded to the targets faster than the native English speakers. With this unexpected finding, a trade-off effect was suspected and examined by running a concordance analysis of accuracy and latency. The direct comparison revealed no systematic trade-off effect. The Chinese students took the longest time to process the given targets in the three shapes. The Chinese participants' processing speed showed greater differences from the natives' performance than that of the Korean counterpart, especially in the low-frequency targets of both normal and alternated fonts. In the comparison of the two L2 groups, significant differences between the Korean and Chinese participants were found across the three font types [$t(19) = -12.47$, $p < .001$ for the normal fonts; $t(19) = -13.47$, $p < .001$ for the alternated fonts; $t(19) = -7.53$, $p < .001$ for the inverse fonts].

The aspect of different performance on the lexical variation of word or nonword stimuli was also compared for accuracy and processing speed. Interestingly, as Fig. 2a shows, all three groups recognized nonwords more accurately than words in the normal fonts, as indicated by the higher accuracy value of nonwords than words. It appeared that deviations from the phonotactic rules expressed in the nonword stimuli were easier to detect than words in the normal fonts. This held constant across the three L1 groups. However, word advantages were found in the manipulated fonts (i.e., alternated and inverse fonts) for the three groups. The Chinese participants had the greatest difficulty with the nonword inverse fonts.[9]

When compared with respect to latency, the data showed a relatively systematic pattern between words and nonwords, in which nonword processing took longer time than word processing across the L1 groups. The inverse fonts took the longest time, followed by the mixed fonts, and normal fonts. A larger disadvantage for nonwords than words was consistent with the results of previous studies (Besner and McCann 1987; Lavidor et al. 2002). Figure 2b

---

[9] Since the magnitude of the effect sizes for the lexical variation was similar to that of word frequency, the effect sizes are not reported here.
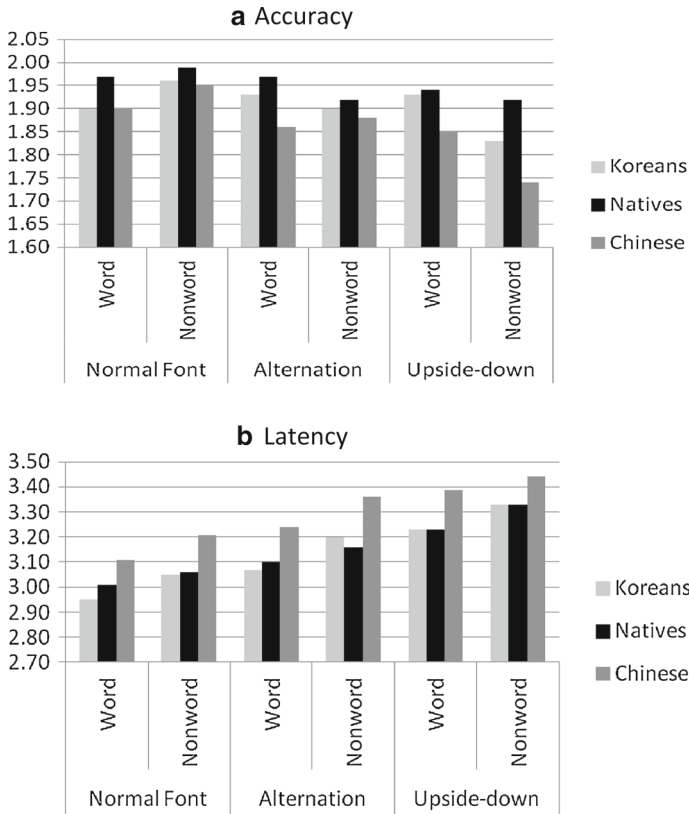
**Fig. 2** Accuracy and latency by lexical variation

displays the three groups' latency data across the fonts by the lexical variation of words and nonwords.

Next, the three groups' performance of accuracy and latency with respect to the number of syllables was compared. Compared with the longer stimuli, overall, accuracy in the shorter target items was higher. Figure 3a shows the three groups' performance on the number of syllables. Interestingly, the Korean and Chinese participants did better on the two-syllable targets than one-syllable ones in the normal fonts. It was speculated that the source of this counter-intuitive result lied in word frequency, given that reading in English is sensitive to word frequency due to the irregular grapheme–phoneme correspondence. The difference of word frequency among one-syllable, two-syllable, and three-syllable word cohorts was computed using the mean of word frequency in each group. The difference between the one-syllable and two-syllable words was 110,469.5, while that between two-syllable and three-syllable words was 11,615.5. Given the great disparity of word frequency in the two groups of word targets, it was assumed that word frequency could be attributable to the Korean- and Chinese-speaking learners' performance on the two-syllable target. The two L2 groups also did better on the four-syllable stimuli than three-syllable stimuli, and generated a very similar pattern in the upside-down shapes.

When it came to the performance on the number of syllables, latencies gradually increased as the number of syllables increased as a function of processing time. This held consistent across the three visual shapes. The native speakers and Korean participants showed a very
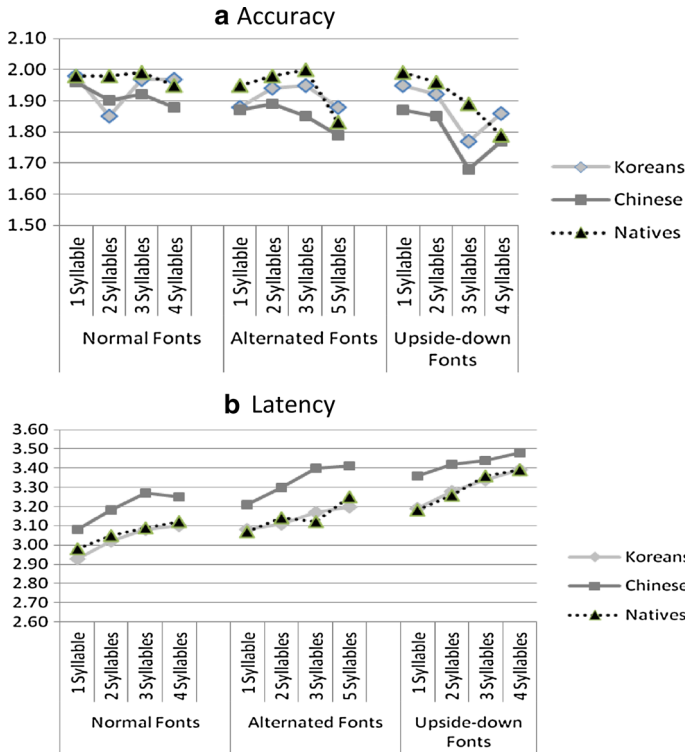
**Fig. 3** Accuracy and latency of number of syllables by the visual shapes

similar trend among the three fonts. The Chinese participants consistently took longer time to process the targets than the Koreans and natives.

In short, the second hypothesis was supported. The Chinese participants were more sensitive to the visual disruptions than the Korean counterpart, resulting in greater noise effects.

### The Role of Word Frequency and Stimulus Length in Word Recognition (Research Question 3)

A set of hierarchical regression analyses was performed by font type in order to gauge the magnitude of variance explained by base-word frequency and stimulus length, using each group's accuracy and latency as dependent variables, and word frequency, the number of letters, and the number of syllables as predictors. The order of entry was determined on the basis of the role played in word recognition found in the literature. For the Korean participants, none of the three predictors explained a significant variance in accuracy across the font types. Concerning the processing speed, word frequency accounted for 46 % of the variance in the normal fonts [$F(1, 18) = 15.01, p < .01$], 38 % of the variance in the alternated fonts [$F(1, 18) = 11.24, p < .01$], and 20 % of the variance in the inverse fonts [$F(1, 18) = 4.48, p < .05$]. The number of syllables explained an additional variance of 11 % in the normal fonts [$F(1, 17) = 4.67, p < .05$] and 30 % in the inverse shapes [$F(1, 17) = 9.93, p < .01$].

For the Chinese participants, word frequency accounted for 25 % of the variance in the normal fonts [$F(1, 18) = 5.96, p < .05$] for accuracy. The number of the letters also

**Table 3** Hierarchical regression analyses for the three groups by the visual shapes

| Dependent variables | Predictors | Normal fonts | | Alternated fonts | | Upside-down fonts | |
|---|---|---|---|---|---|---|---|
| | | $R^2$ | $\triangle R^2$ | $R^2$ | $\triangle R^2$ | $R^2$ | $\triangle R^2$ |
| *Koreans* | | | | | | | |
| Accuracy | Step 1. Word frequency | .09 | – | .08 | – | .20 | – |
| | Step 2. Number of syllables | .09 | .00 | .13 | .05 | .25 | .05 |
| | Step 3. Number of letters | .09 | .00 | .30 | .17 | .26 | .01 |
| Latency | Step 1. Word frequency | .46** | – | .38** | – | .20* | – |
| | Step 2. Number of syllables | .57*** | .11* | .40** | .02 | .50** | .30** |
| | Step 3. Number of letters | .63*** | .06 | .43** | .03 | .50** | .00 |
| *Chinese* | | | | | | | |
| Accuracy | Step 1. Word frequency | .25* | – | .14 | – | .17 | – |
| | Step 2. Number of syllables | .25* | .00 | .16 | .02 | .36* | .19* |
| | Step 3. Number of letters | .28* | .03 | .16 | .00 | .36* | .00 |
| Latency | Step 1. Word frequency | .49** | – | .43** | – | .15 | – |
| | Step 2. Number of syllables | .57** | .08 | .54** | .11* | .54** | .39** |
| | Step 3. Number of letters | .58** | .01 | .55** | .01 | .55** | .01 |
| *Native speakers* | | | | | | | |
| Accuracy | Step 1. Word frequency | .15 | – | .30* | – | .50*** | – |
| | Step 2. Number of syllables | .15 | .00 | .30* | .00 | .56*** | .16* |
| | Step 3. Number of letters | .23 | .08 | .34* | .04 | .60** | .00 |
| Latency | Step 1. Word frequency | .60** | – | .42** | – | .40** | – |
| | Step 2. Number of syllables | .66** | .06 | .55 | .13* | .62*** | .22* |
| | Step 3. Number of letters | .68** | .02 | .55 | .00 | .62*** | .00 |

\* $p < .05$; \*\* $p < .01$; \*\*\* $p < .001$

explained a unique variance of 19 % in the inverse fonts [$F(1, 17) = 4.81 p < .05$] for accuracy. As for latency, word frequency explained 49 % of the variance in the normal fonts [$F(1, 18) = 17.35, p < .01$] and 43 % of the variance in the alternated fonts [$F(1, 18) = 13.58, p < .01$]. The number of syllables accounted for an additional variance of 11 % in the alternated fonts [$F(1, 17) = 5.27, p < .05$] and 39 % in the inverse fonts [$F(1, 17) = 14.44, p < .01$].

For the native speakers, word frequency was predictive of the alternated fonts [30 % of the variance; $F(1, 18) = 7.75, p < .05$] and the upside-down fonts [50 % of the variance, $F(1, 18) = 18.14, p < .001$] for accuracy. As far as latency was concerned, word frequency explained a significant variance in the three fonts: 60 %; [$F(1, 18) = 26.54, p < .001$] for the normal fonts; 42 %, $F(1, 18) = 12.89, p < .01$ for the alternated fonts; 40 %, $F(1, 18) = 11.95, p < .01$) for the upside-down fonts. The number of syllables accounted for an additional variance in the alternated fonts [13 %, $F(1, 17) = 4.93, p < .01$] and a unique variance in the inverse fonts [22 %, $F(1, 17) = 9.52, p < .01$]. Table 3 shows the results of a series of hierarchical regression analyses of word frequency, the number of letters, and the number of syllables by font and language group.

In short, the third hypothesis was supported. The Chinese participants showed greater holistic noise effects than the Korean counterpart.

## Discussion

The performance of Korean and Chinese speakers of English as L2 on a lexical decision test was compared to that of native English speakers in terms of accuracy and latency. The two L2 groups were included in this study because their L1s were different from each other with respect to the depth of orthography and grain sizes involved in word reading. The performance of the Korean and Chinese participants was considered comparable because (1) the base words of the stimuli was at the grade equivalency of 6.7, (2) all the participants were college or graduate students who have already passed the threshold of isolated word reading and gained the automaticity of word reading, and (3) importantly, the accuracy of normal-font word reading, which was considered a baseline datum, did not show a significant difference between those two groups.

The alternated font was utilized to examine intraword noise effects (i.e., sublexical visual information is disrupted within the lexicon), while the inverse shape was used to look at holistic noise effects (i.e., sublexical information is preserved but the whole lexical unit is inverted). Overall, the results were compatible with those of previous studies which supported that various factors, such as the L1 system, word frequency, lexical variation, and lexical length, play a role in the recognition of written words in L2 (Akamatsu 1999, 2003; Allen et al. 1995; Besner 1983; Besner and McCann 1987; Lavidor et al. 2002; Reingold et al. 2010).

There were main and interaction effects in accuracy, indicating that the three L1 groups performed differently and that their L1s affected their performance on the three font types. As expected, the native speakers were least affected by the visual noise among the three language groups. As Ziegler and Goswami's (2005) psycholinguistic grain size theory (i.e., each language has a minimal processing unit, such as phoneme, onset-rime, body-coda, or syllable) suggests, English speakers seem to apply more flexible grain sizes (i.e., more than one grain size) in word recognition. It is because a reliance on small grain sizes is not sufficient enough to read English words correctly due, in part, to the inconsistent grapheme–phoneme correspondence. Consequently, the native English speakers were less sensitive to lexical noise than the speakers of other languages above and beyond their L1 status. With respect to latency, there was no significant difference between the native speakers and the Korean participants, whereas the native speakers performed differently from the Chinese participants. No significant difference between the natives and Koreans might have stemmed from the shared alphabetic characteristics in English and Korean. The findings supported, by and large, the hypothesis that the L1 difference (i.e., phoneme-based language vs. syllable-based language) would produce a different pattern in L2 English word recognition by the Korean- and Chinese-speaking participants. The differences in the orthographic depth and grain sizes in L1s appeared to affect the resolution of noisy stimuli in English as FL and L2, as indicated by the effects of intraword noise and holistic noise.

The two L2 groups showed a similar pattern of nonsignificant correlations between accuracy and latency; that is, no significant association was found between accuracy and latency in the Korean and Chinese participants, except the significant correlation in the inverse fonts for the Chinese participants. The native speakers showed significant correlations between accuracy and latency across the three font types. This finding indicates that the two L2 groups' accuracy scores on a lexical decision test are not subject to processing speed in L2 word recognition.

In reading normal cases, the Korean- and Chinese-speaking participants showed a very similar pattern in high and low frequency target stimuli. However, the Chinese participants were more sensitive to word frequency in the alternated and inverse fonts than in the normal fonts. The greatest disparity was shown in the inverse fonts for accuracy and in the alternated

fonts for latency, when compared to the performance of the native speakers. Since the Chinese participants are accustomed to whole-word processing due to the syllabic and logographic features of their L1, lexical noise seems to affect the Chinese-speaking participants more than the Korean counterpart.

The lexical variation (i.e., words or nonwords) produced similarities and differences in the performance of the Korean and Chinese participants. Compared to the native speakers' performance, the two L2 groups showed striking similarities in both words and nonwords of the normal fonts. A phonotactic advantage is reduced in nonwords than real-words; as a result, the violation of spelling conventions would have an adverse effect. As Allen et al. (1995) have noted, familiar words have more configurational features than nonwords. The Chinese-speaking participants were more affected by the alternated fonts in the real-word targets than the nonword stimuli. This result is consistent with the finding of previous studies that reported stronger case alternation effects of real words than nonwords in a lexical decision task (Besner 1983; Lavidor et al. 2002). The familiarity discrimination mechanism involved in word recognition seems to be a plausible explanation for the stronger effects of intraword noise on real words than nonwords (Besner 1983; Lavidor et al. 2002). Since case alternation disrupts the conventional word shape, the intraword noise created by the alternated fonts interrupts both automaticity of word reading and efficient use of the mechanism of familiarity discrimination. According to the notion of the familiarity discrimination mechanism, case alternation yields less disruptive effects on nonword targets than real words, because nonwords are less familiar than words. Regarding the inverse fonts, the nonword targets produced the biggest disparity between the native speakers and the two L2 groups in accuracy. Reduced cuing information in the inverse shape might have made familiarity less influential, as suggested by the L2 groups' performance on nonwords. As indicated earlier, the Chinese participants seem to be more affected by the visual distortion than the Korean counterpart. This implies that the L1 feature influences L2 word recognition (Akamatsu 1999, 2003; McBride-Chang et al. 2008).

The results of hierarchical regression analyses demonstrated that word frequency accounted for the significant variance in the accuracy score of word recognition of distorted fonts in the native English speakers, whereas the two L2 groups were less affected by word frequency. This finding is consistent with the results of previous research that English speakers are sensitive to word frequency (Allen et al. 1995). Processing speed was influenced by both word frequency and target length across the three language groups.

In a nutshell, the most important finding in this study centers on the greater influence of visual distortion shown by the Chinese nonnative speakers of English than the Korean counterpart. The implication of this finding is twofold: (1) visual recognition differences and (2) grain size differences. First, the difference of visual recognition between Korean and Chinese mainly stems from the linguistic distance between L1 and L2. Given that the linguistic distance between Korean and English is closer than that between Chinese and English, it is congruent with the hypothesis that the Chinese participants would show more sensitive reactions than the Korean counterpart to visually distorted L2 English stimuli. The visual recognition difference is also related to L1 orthographic effect.[10] Second, the difference

---

[10] A possibility of L2 word-recognition proficiency effects was suspected, given the seemingly different performance between the Korean and Chinese participants. In order to address this, the possibility was examined in three ways: First, the performance on the normal font was treated as a baseline datum. A *t* test revealed no significant difference ($p > .05$) in accuracy between the two groups. Secondly, we found that the grade equivalency of the base words equates to grade 6.7, which means that students in the first year of middle school are expected to know all the base words used in this study. Lastly, since the participants were college or graduate students and all of them had passing scores of TOEFL, it was assumed that

of the grain size among the three languages points to a possibility that the smaller grain size L1 has, the less susceptible to visual disruption in L2. As stated earlier, the Korean language as alphasyllabary has a smaller grain size than Chinese. The utility of a smaller grain size seems to play a part in the resolution of intraword noise and holistic noise.

From a microscopic view, the differential effect of the two types of visual distortions can be explained on the basis of the greater holistic noise effect, along with the inverse font's stronger correlation with the word length as well as a greater variance than the alternated fonts explained by the number of syllables. One explanation would be L1 orthographic effects (Akamatsu 1999, 2003). Due to the linguistic difference, Chinese speakers tend to process visual stimuli holistically, whereas Korean speakers are prone to process visual words atomically. Hence, the finding that the Chinese participants show greater sensitivity to visual noise than the Korean counterpart is congruent with the expectation. What is more interesting in the finding of this study is the greater degree to which the inverse fonts affect Chinese speakers' word recognition than the alternated font. The greater holistic noise than intraword noise may also point to the familiarity effect. Lower and upper cases in English are more common in English than inverted fonts in prints. Besides, since Chinese word decoding skills are acquired through practices and drills emphasizing rote memorization (Louie and Louie 2002), it is speculated that the initial acquisition mode of Chinese characters may also have an influence on later L2 word decoding. Further research is warranted to address this issue.

Although this study was built on previous research, the scope of this study makes it unique. Specifically, this study addressed the performance of three language groups, focusing on the two L2 groups with the comparison group of native speakers, as well as the issue of independent contributions of word frequency and word length to the processing of normal fonts and lexically noisy fonts, which is rare in the literature of L2 psycholinguistic research. In addition, the depth and breadth of analyses cover a wide range of lexical intricacies in the understanding of English word processing from multiple angles. This study contributes to the word recognition area in that the findings not only demonstrate the direction of relationship and the magnitude of influences of the different font types and lexical variations on word recognition, but also explain L1 effects of different linguistic systems on L2 English word identification.

To summarize, the Korean-speaking learners were less susceptible to lexical noise, while the Chinese-speaking participants were more sensitive to it. This result suggests that the psycholinguistic grain size in L1 (i.e., phonemes vs. syllables) plays a salient role in word recognition in L2. Given that the configurational quality of lexicons is maintained in words more than in nonwords, due to the phonotactic rules, the mechanism of familiarity discrimination seems to be in operation in the face of noise resolution in word reading.

In order to corroborate the findings of this study as well as to expand the body of knowledge, further research is warranted. First, although the sample size of this study was sufficient to address the research questions, a study with a larger sample size would allow for examinations of different inquiries, including $n$-gram modeling. An investigation of the likelihood that letter strings affect the reader's word recognition by modeling bigrams, trigrams, or $n$-grams in sequences of graphemes would help better understand the role of lexical quality in word recognition. Second, the phonotactic rule or legality of letter strings was not controlled in this study. Future studies that control for specific phonotactic constraints in the stimulus targets would validate the findings of this study. Next, because the sole purpose of this study

---

footnote10 continued

automaticity of word reading was gained. In addition, an additional regression analysis was run controlling for the score of the normal-font accuracy. The results showed no significant difference. Therefore, that possibility was ruled out. A follow-up study is warranted to confirm this.

centered on visual processing, phonological and semantic activation in word recognition was not considered. Studies that have different focal points taking phonology and semantics into consideration would expand the understanding of the extent to which different fonts or shapes play a role in lexical judgment. Last, this study did not investigate the role of cuing information in upper-part-absent and bottom-part-absent stimuli. Future studies that include those two types of targets in a lexical decision task would substantiate the explanation of intraword and holistic noise effects, which are addressed first in this paper.

## References

Adams, M. J. (1990). *Beginning to read*. Cambridge, MA: MIT Press.

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

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

Allen, P. A., Wallace, B., & Weber, T. A. (1995). Influence of case type, word frequency, and exposure duration on visual word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 914–935.

Besner, D. (1983). Basic decoding components in reading: Two dissociable feature extraction processes. *Canadian Journal of Psychology*, *37*, 429–438.

Besner, D., & McCann, R. S. (1987). Word frequency and pattern distortion in visual word identification and production: An examination of four classes of models. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 201–219). Mahwah, NJ: Lawrence Erlbaum.

Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum.

Chen, H.-C., & Vaid, J. (2007). Word frequency modulates the basic orthographic syllabic structure (BOSS) effect in English polysyllable word recognition. *Language and Cognitive Processes*, *22*, 58–82.

Cho, J.-R., & McBride-Chang, C. (2005). Levels of phonological awareness in Korean and English: A 1-year longitudinal study. *Journal of Educational Psychology*, *97*, 564–571.

Davies, M. (2013). The corpus of contemporary American English (COCA). Retrieved February 15, 2012, http://corpus.byu.edu/coca/.

Durgunoglu, A. Y., & Oney, B. (1999). A cross-linguistic comparison of phonological awareness and word recognition. *Reading and Writing*, *11*, 281–299.

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

Frost, R., Katz, L., & Bentin, S. (1987). Strategies for visual word recognition and orthographic depth: A multilingual comparison. *Journal of Experimental Psychology: Human Perception and Performance*, *13*, 104–115.

Goswami, U., Ziegler, J. C., Dalton, L., & Schneider, W. (2003). Nonword reading across orthographies: How flexible is the choice of reading units? *Applied Psycholinguistics*, *24*, 235–247.

Gottardo, A., Yan, B., Siegel, L. S., & Wade-Woolley, L. (2001). Factors related to English reading performance in children with Chinese as a first language: More evidence of cross-language transfer of phonological processing. *Journal of Educational Psychology*, *93*, 530–542.

Kintsch, W. (1994). Text comprehension, memory, and learning. *American Psychologist*, *49*, 294–303.

Lavidor, M., Ellis, A. W., & Pansky, A. (2002). Case alternation and length effects in lateralized word recognition: Studies of English and Hebrew. *Brain and Cognition*, *50*, 257–271.

Lee, C. (1999). A locus of the word-length effect on word recognition. *Reading Psychology*, *20*, 129–150.

Louie, B. Y., & Louie, D. H. (2002). Children's literature in the People's Republic of China: Its purposes and genres. In W. Li, J. S. Gaffney, & J. L. Packard (Eds.), *Chinese children's reading acquisition: Theoretical and pedagogical issues* (pp. 175–193). Boston, MA: Kluwer.

McBride-Chang, C., Tong, X., Shu, H., Wong, A. M.-Y., Leung, K.-W., & Tardif, T. (2008). Syllable, phoneme, and tone: Psycholinguistic units in early Chinese and English word recognition. *Scientific Studies of Reading*, *12*, 171–194.

McBride-Chang, C., Bialystok, E., Chong, K., & Li, Y. P. (2004). Levels of phonological awareness in three cultures. *Journal of Experimental Child Psychology*, *89*, 93–111.

Pae, H. K. (2011). Is Korean a syllabic alphabet or an alphabetic syllabary? *Writing Systems Research*, *3*, 103–115.

Pae, H. K., Greenberg, D., & Williams, R. S. (2012). An analysis of differential error patterns on the Peabody Picture Vocabulary Test-III in children and struggling African-American adults readers. *Reading and Writing: An Interdisciplinary Journal*, 25, 1239–1258.

Pae, H. K., Sevcik, R. A., & Morris, R. D. (2010). Cross-language correlates in phonological process and naming speed: Evidence from deep and shallow orthographies. *The Journal of Research in Reading*, 33, 335–436.

Perea, M., Rosa, E., & Gomez, C. (2002). Is the go/no-go lexical decision task an alternative to the yes/no lexical decision task? *Memory and Cognition*, 30, 34–45.

Perfetti, C. A., & Liu, Y. (2005). Orthography to phonology and meaning: Comparisons across and within writing systems. *Reading and Writing: An Interdisciplinary Journal*, 18, 193–210.

Reingold, E. M., Yang, J., & Rayner, K. (2010). The time course of word frequency and case alternation effects on fixation times in reading: Evidence for lexical control of eye movements. *Journal of Experimental Psychology*, 36, 1677–1655.

Seidenberg, M. S. (2012). Writing systems: Not optimal, but good enough. *Behavioral and Brain Sciences*, 35, 305–307.

Simpson, G. B., & Kang, H. (2004). Syllable processing in alphabetic Korean. *Reading and Writing: An Interdisciplinary Journal*, 17, 137–151.

Wang, M., Koda, K., & Perfetti, C. A. (2003). Alphabetic and nonalphabetic L1 effects in English word identification: A comparison of Korean and Chinese English L2 learners. *Cognition*, 87, 129–149.

Wang, M., & Cheng, C. (2008). Subsyllabic unit preference in young Chinese children. *Applied Psycholinguistics*, 29, 291–314.

Weaver, C. (1988). *Reading process and practice: From socio-psycholinguistics to whole language* (2nd ed.). Portsmouth, NH: Heinemann.

Woodcock, R. W. (1998a). *Woodcock reading mastery tests-NU*. Circle Pines, MN: American Guidance Service.

Woodcock, R. W. (1998b). *Woodcock reading master tests-NU: Examiner's manual*. Circle Pines, MN: American Guidance Service.

Yoon, H. K., Bolger, D. J., Kwon, O. S., & Perfetti, C. A. (2002). Subsyllabic units in reading: A difference between Korean and English. In L. Verhoeven, C. Elbro, & P. Reitsma (Eds.), *Precursors of functional literacy* (pp. 139–163). Amsterdam: John Benjamins Publishing Co.

Ziegler, J., & Goswami, C. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin*, 131, 3–29.