ABSTRACT

Title of Dissertation:	LEXICAL DEVELOPMENT AND MASKED ORTHOGRAPHIC PRIMING IN THE SECOND LANGUAGE
	Kichan Park, Doctor of Philosophy, 2022
Dissertation Directed by	Professor, Kira Gor, Second Language Acquisition

The fuzzy lexical representations (FLR) hypothesis proposes that form encoding of words in a second language (L2) is often fuzzy, and this concerns both phonological and orthographic representations. FLR occur because of difficulties in encoding of L2 word forms as well as insufficient L2 experience. The FLR hypothesis also suggests that fuzzy L2 orthographic representations are the reason for the weak lexical competition for orthographic neighbor prime-target pairs in the L2 that has been observed in previous research (e.g., Jiang, 2021). However, this hypothesis also assumes that as orthographic representations become robust along with learners' L2 experience, L2 words are eventually able to take part in lexical competition just like first language (L1) words. The current study tests these hypotheses using the individual-differences measures of the quality (orthographic precision) and the quantity (vocabulary size) of orthographic representations. At the same time, this study explores the relationship between sound perception (word and phoneme identification) of nonnative contrasts (e.g., the /l/-/ɹ/

contrast for Korean L2 learners of English), phonolexical encoding, and form facilitation for minimal pairs with these contrasts.

A masked priming LDT was employed, in which minimal pairs with a nonnative phonological contrast (e.g., *read-LEAD*) and minimal pairs without a confusing phonological contrast (e.g., *dear-TEAR*) were used as the prime and target. Before the experiment, it was predicted that low-proficiency L2 speakers would show significant form facilitation under all prime conditions. On the other hand, medium-proficiency L2 speakers were expected to show evidence of emerging lexical competition (a null priming effect) for prime-target pairs without a difficult phonological contrast (e.g., *dear-TEAR*), although they would still show form facilitation for minimal pairs with a nonnative phonological contrast (e.g., *read-LEAD*). The facilitation for the latter pairs was predicted to occur because of less successful orthographic encoding of these pairs caused by fuzzy phonological representations of L2 words with difficult phonological contrasts. It was further expected that high-proficiency L2 speakers would show a nativelike pattern of form priming across all the prime conditions.

Thirty L1 speakers and 90 L2 learners of English with a wide range of L2 proficiency were recruited for the experiment. In auditory word and phoneme identification tasks, L2 speakers showed less accurate identification of the /l/-/ɹ/ contrast compared to L1 speakers indicating that they indeed had problems in accurate sound perception and/or phonological categorization of the nonnative contrast as had been predicted.

In the masked priming LDT, L1 speakers showed a null priming effect across the prime conditions. L2 speakers showed significant form priming for words with the /l/-/ɹ/ contrast but not for other words without a difficult contrast. When form priming in each L2 participant group was examined separately, low- and medium-proficiency L2 speakers showed significant facilitation for pairs with the /l/-/ɹ/ contrast, but high-proficiency L2 speakers showed a null priming effect for these pairs as L1 speakers did. This finding supports the prediction of the current study. At the same time, the influence of global proficiency, as measured by a cloze test, on the orthographic form priming was statistically non-significant. Furthermore, form facilitation for prime-target pairs without a confusing contrast (e.g., *dear-TEAR*) was not significant even in low-proficiency L2 participant groups.

Through a series of investigations on the relationships between the form priming found in L2 speakers and their performance on individual-differences measures (spelling, vocabulary, word identification and phoneme identification tasks), the present study discovered that form facilitation was significantly modulated by L2 speakers' orthographic precision (spelling scores). Moreover, it was found that the influence of orthographic precision on the form facilitation was more prominent for words that were more difficult for accurate phonological encoding, and as a consequence, orthographic encoding (i.e., minimal pairs with the /l/-/ɹ/ contrast) than others without a confusing contrast. These findings support the FLR hypothesis which argues for the role of the quality of orthographic representations in lexical competition between orthographic neighbors. The role of vocabulary size (vocabulary scores) was also found for four-letter stimuli indicating that the development of the size of the mental lexicon also affects lexical competition. On the other hand, no modulating role was observed of accurate word or phoneme identification of nonnative contrasts in form priming for minimal pairs with these contrasts.

Based on these findings, this study suggests that (1) the orthographic form facilitation discovered at initial stages of L2 lexical development is due to fuzzy L2 orthographic representations. In addition, it claims that (2) as L2 speakers establish a larger and more precise L2 lexicon, L2 words can take part in lexical competition just as L1 words do. It also proposes that (3) the establishment of precise orthographic (or phonological) representations of L2 words with a confusing phonological contrast is more challenging than those without a difficult contrast. (4) Finally, although the observed weak effect of sound perception on form priming seems to indicate no systematic relationship between the development of phonological categorization ability and the form facilitation for these words, the present study contends that it may be premature to draw a conclusion about the role of phonolexical representations involving a nonnative contrast in orthographic representations. Indeed, the results may be due to methodological limitations of the word and phoneme identification tasks as a measure of the quality of phonological representations.

LEXICAL DEVELOPMENT AND MASKED ORTHOGRAPHIC PRIMING
IN THE SECOND LANGUAGE

by

Kichan Park

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2022

Advisory Committee:
Professor Kira Gor, Chair
Professor Nan Jiang
Associate Professor Donald Bolger
Associate Graduate Faculty Stefanie Kuchinsky
Professor Min Wang, Dean's Representative

Acknowledgements

It is an unbelievable feeling of satisfaction to have completed this dissertation, and I owe this feeling to the help and support of so many amazing people.

First of all, I would like to express my deepest appreciation to my advisor, Dr. Kira Gor. Whenever I was stranded in the middle of my PhD study, she always helped me get through numerous challenges and consistently supported me to find a new route to successfully complete my study. Her philosophy and perspectives on second language vocabulary processing motivated the current study, and her profound insights about experimental design and data interpretation considerably enhanced the quality of the present study.

I am also deeply thankful to Dr. Nan Jiang who guided me to acquire the solid background required for conducting psycholinguistic experiments. All his careful instruction and advice deepened my understanding of second language research and became the firm basis of the current study. As a committee member, he also provided me with invaluable new insights about the topic of this study based on many years of deep reflection.

My gratitude extends to my other committee members. Dr. Donald Bolger, Dr. Min Wang and Dr. Stefanie Kuchinsky all eagerly agreed to serve as committee members for this study and provided precious constructive feedback that went far beyond their responsibilities. Their great care about and attention to the present study have been transformational for this dissertation.

I am also very grateful to other professors in the Second Language Acquisition program. Although he could not attend the final defense for my dissertation, Dr. Steven Ross served as a committee member for the proposal defense and offered me good advice for proper statistical analysis. His instruction on research methods also helped me correctly understand previous research and correctly design my own study. Dr. Mike Long and Dr. Robert DeKeyser were always so passionate in their courses and assisted me in building my own philosophy of second language acquisition. Without having had a chance to learn from them, this dissertation would not have been possible.

I would like to thank many current students in and alumni from the Second Language Acquisition program as well. Stefanie Simmons and Jonathan Malone aided me with materials development, and So-Hye Park, Dr. Jiyong Lee and Dr. Hyojin Jeong assisted with participant recruitment. Many of my friends and colleagues took part in a pilot test for the dissertation study and offered ceaseless emotional encouragement during my study. All of their support was of big help for me.

Finally, I am deeply grateful to the Korean Navy that offered me a precious chance to study in one of the most prestigious Second Language Acquisition programs in the world at the University of Maryland. They financially supported my PhD study and allowed me to concentrate on it for five years.

And last, but not least, I am so very grateful to have the three people in my life, who made all the difference in the world – my wife Yunju Yang and my children Junseo and Jane. I do not think I will ever be able to express how much your love and support has meant to me throughout those five trying years of our life. Your support and devotion picked me up and pushed me when I was losing faith in myself and could not go any further; you are my anchor and my safe place. Thank you for believing in me and being there for me at every step of the way.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

L1: First language

L2: Second language

FLR: Fuzzy lexical representations

LDT: Lexical decision task

PLE: Prime lexicality effect

RT: Reaction Time

SD: Standard deviation

SOA: Stimulus onset asynchrony

TOEIC: Test of English for International Communication

Chapter 1. Introduction

Much psycholinguistic research has been conducted to investigate the organization and characteristics of the mental lexicon, a hypothetical storage of words in long-term memory where the formal (phonological and orthographic), morphological, syntactic and semantic information of words is encoded (Gor, forthcoming; Jiang, 2000). This body of research is based upon the widely agreed assumption that words represented in the mental lexicon are not just stably stockpiled but interact with one another. Even though many studies have revealed how similar first language (L1) word forms interact, it is only recently that interactions between second language (L2) word forms have begun to be examined (see below). The finding that L2 words appear to interact differently from L1 words during spoken word processing has prompted the *fuzzy lexical representation (FLR) Hypothesis* (Cook & Gor, 2015; Gor, 2018, forthcoming; Gor & Cook, 2020; Gor et al., 2021).

On the other hand, distinctive patterns of interactions between L2 words have also been reported in another body of research on written word processing. Although some accounts similar to the FLR hypothesis have been suggested to explain atypical interactions observed in L1 lexicon (e.g., Andrews & Hersch, 2010; Castles, Davis, Cavalot, & Forster, 2007; Perfetti, 2007; see Section 2.1.1 below), the applicability of these accounts to the L2 lexicon has not yet been fully examined. Instead, a hypothesis assuming separate L1 and L2 lexicons was proposed to explain non-nativelike behaviors of L2 words (Qiao & Forster, 2017). Given this context, the main purpose of the present study was to explore whether the FLR hypothesis can explain why L2 words are processed differently from L1 words during visual word recognition. Before reviewing relevant research on visual word identification, this chapter first explains the concept of the FLR hypothesis and introduces the findings of studies that support this hypothesis. It then briefly addresses key assumptions that the FLR hypothesis makes in consideration of written word recognition. These assumptions guide the following chapters.

1.1. The Fuzzy Lexical Representations Hypothesis

1.1.1. The Concept of the Fuzzy Lexical Representations Hypothesis

One of the phenomena that has been used for investigation of lexical representations is phonological priming occurring during auditory word recognition. In this method, participants are asked to respond to a target that is preceded either by a similarly sounding or an unrelated prime. Slowiaczek and Hamburger (Experiments 1 and 2, 1992), for example, found that when native English speakers were asked to perform a phonologically primed, single-word shadowing task (a task in which participants are asked to repeat the target word that they heard), their responses to the word target (e.g., *still*) were facilitated if the word prime shared the first phoneme with the target (e.g., *smoke*) compared to no shared phonemes with the target (e.g., *dream*). However, this facilitation induced by the phonological prime was not observed if the prime shared the initial three phonemes with the target (e.g., *stiff*). When this type of prime was presented, participants' response times were significantly slower than when word primes with an initial one-phoneme overlap were used. They also appeared slower relative to when unrelated primes preceded the target[1]. At the same time, Slowiaczek and Hamburger discovered that if the

[1] Slowiaczek and Hamburger (1992) report 2 ms of numerical inhibition in Experiment 1A (with auditory primes), 18 ms of numerical inhibition in Experiment 1B (with visual primes), 16 ms of numerical inhibition in Experiment 2A (with auditory primes) and 27 ms of numerical inhibition in Experiment 2B (with visual primes) if the inhibition is defined as the difference between the response latencies to the target after encountering a related prime (i.e., the prime that shared the initial three phonemes with the target) and those after encountering an unrelated prime (i.e., the prime that shared no phonemes with the target). However, the researchers did not report the results of statistical analysis for these inhibitory priming effects because they were more interested in

prime was a nonword (Experiment 3A), whether the prime shared initial one phoneme or three phonemes with the word target did not matter. In that case, a facilitative priming effect from a one phoneme overlap was not significantly different from that from a three-phoneme overlap.

Slowiaczek and Hamburger (1992) interpreted that the facilitative effect from an initial one-phoneme overlap between the (word or nonword) prime and the target occurs prelexically and originates from activation of a sub-lexical component (i.e., a phoneme) of the target by the prime. They then thought that the inhibition[2] from word primes with an initial three-phoneme overlap occurs lexically and is the outcome of competition for selection between lexical units that have similar phonological forms. In other words, they assumed that when a prime that shares its initial three phonemes with other words is provided, it triggers lexical competition because this overlap is sufficient to activate candidates for selection (Marslen-Wilson, 1987). Slowiaczek and Hamburger (1992) also considered that the prime lexicality effect (PLE), which refers to the different patterns of priming from word primes (inhibition) and nonword primes (a null effect)[3], with an initial three-phoneme overlap is also evidence for the assumption that competition occurs only between lexical units (i.e., the word prime and the word target).

---

comparing response latency to the target after encountering a prime with an initial one-phoneme overlap with that after encountering another prime with an initial three-phoneme overlap.

[2] In Slowiaczek and Hamburger's study (1992), facilitation refers to faster reaction times from an initial one-phoneme overlap between the prime and the target compared to those from zero-phoneme overlap. On the other hand, inhibition refers to slower reaction times from a three-phoneme overlap compared to those from a one-phoneme overlap. Therefore, the way Slowiaczek and Hamburger operationalized inhibition is different from that in other studies in which response times to a target preceded by a related prime are compared to response times to a target preceded by an unrelated prime.

[3] In Slowiaczek and Hamburger's study (1992), the null effect from nonword phonological primes was also operationalized by comparing the priming effects from a one-phoneme overlap (e.g., /bIs/-*blood*) and a three-phoneme overlap (e.g., /blʌn/-*blood*). As noted in Footnote 1, these researchers did not report the results of statistical analysis for comparison of the priming from a three-phoneme overlap with that from a zero-phoneme overlap (e.g., /græks/-*blood*).

In a subsequent study (Dufour & Peereman, 2003b), a similar pattern of phonological priming was observed in native speakers. In a shadowing task, French L1 speakers responded significantly more slowly to the word target when it was preceded by a word prime that shared its initial three phonemes with the target compared to when it was preceded by an unrelated prime (in Experiments 1A and 1B)[4]. Similarly, Gor and Cook (2020) used a lexical decision task (LDT) and found that L1 Russian speakers' lexical decisions were inhibited when the target followed primes that overlapped with the target in its initial three phonemes relative to when it followed unrelated primes. However, interestingly, L2 speakers did not show the same pattern. Although their lexical decisions were also inhibited for high-frequency prime-target pairs, low-frequency stimulus pairs produced facilitation in L2 speakers unlike in L1 speakers. Gor and Cook (2020) explained their discovery based on the FLR hypothesis.

The FLR hypothesis (Cook & Gor, 2015; Gor, 2018, forthcoming; Gor & Cook, 2020; Gor et al., forthcoming) refers to the idea that when lexical representations are fuzzy, they do not function as robust lexical representations do. In a broader sense, fuzziness means inexact or ambiguous registrations in memory of the different components or dimensions of the lexical representation (i.e., the dimension of linguistic domains such as orthography, phonology and semantics, the dimension of mappings and the dimension of network). This concept may explain many distinctive phenomena observed during L2 lexical processing (see the Ontogenesis Model of the L2 lexical representation, Bordag, Gor, & Opitz, 2021). The FLR hypothesis does not argue that atypical behaviors observed during lexical processing are limited only to L2 words,

[4] Dufour and Peereman (2003b) also showed that a three-phoneme overlap is not always a sufficient condition for lexical competition. Specifically speaking, they demonstrated that the strength of lexical competition is determined not only by the amount of overlap in initial phonemes but also by the number of mismatches in the final phonemes (Experiments 2, 3 and 4).

and that they can also occur during some L1 word processing if their lexical representations are underdeveloped. However, it posits that while most L1 representations are properly encoded and fully specified, many L2 lexical representations do not reach this stage due to cognitive limitations in encoding L2 word forms and form-meaning mappings (Gor et al., 2021).

Although the scope of discussion on fuzzy representations can be very wide as noted above, the current discussion focuses on the fuzziness present in the domain of phonological representations. Gor and Cook (2020) argue that form facilitation for L2 words observed in their phonologically primed auditory LDT occurred because the phonological forms of low-frequency L2 words that are represented in L2 speakers' memory are fuzzy. For that reason, it was considered that lexical competition, which requires the prime and target both with robust phonological representations, could not take place. If either the prime or the target is weak, the prime would activate only sub-lexical components (i.e., overlapped phonemes) of the target but would not trigger lexical competition, thus leading to facilitation given the absence of strong lateral inhibition between lexical units.

1.1.2. The Causes of Fuzzy L2 Phonological Representations

The occurrence of fuzzy L2 phonological representations has been attributed to unfaithful phonological encoding as well as insufficient exposure to L2 words (Gor, 2018, forthcoming). To be more specific, L2 learners often find it difficult to discriminate two words with nonnative phonological contrasts. For instance, Pallier, Colomé and Sebastián-Gallés (2001) showed that L2 speakers could not differentiate two different words that had confusing phonological contrasts. In Catalan, the phonemes /e/-/ɛ/, /o/-/ɔ/ and /s/-/z/ are contrasted whereas they are not in Spanish because Spanish has five vowels (/a/, /e/, /i/, /o/ and /u/) and only voiced fricatives (but not unvoiced fricatives). This inventory of Spanish vowels and consonants differs from that of Catalan that includes eight vowels (/a/, /e/, /ɛ/, /i/, /o/, /ɔ/, /u/ and /ə/) and both voiced and unvoiced fricatives. Pallier et al. (2001) used the medium-term repetition priming method, in which the prime and the target were presented with an interval of 8~20 other stimuli while participants were making lexical decisions on both the prime and the target. When each member of minimal pairs with difficult phonological contrasts in Catalan were used as the word prime and the word target (e.g., /nɛtə/-/netə/), Spanish-dominant bilinguals responded faster to the target since a repetition priming effect from the perceptually same phonological form occurred for the minimal pairs, whereas this facilitation was not detected in Catalan-dominant bilinguals. This finding shows that L2 speakers lack the ability to differentiate non-native phonological contrasts. Based on this finding, Pallier and his colleagues (2001) suggested that "if listeners have difficulties perceiving an L2 phonemic contrast, they will represent L2 word pairs with that contrast as homophones" (p. 448).

The difficulty in accurate phonological encoding of L2 words with confusing phonological contrasts was also observed in a study that used the eye-tracking method. In Dutch, the phonemic contrasts /æ/-/ɛ/ and /aɪ/-/eɪ/ that are both used in English are not discriminated since Dutch has a vowel labeled /e/ and a diphthong labeled /ɛɪ/, but no /æ/ or /aɪ/. In Weber and Cutler's study (2004), L1 English speakers and Dutch L2 learners of English were asked to choose a picture that corresponded to an aurally presented English word out of four candidates, and one of the four pictures (the distractor picture) represented another word that contained a vowel that L1 Dutch speakers were likely to confuse with the vowel in the target picture name (e.g., *panda-pencil*). While participants were performing this task, their eye-movements were tracked. Under this experimental condition, L2 speakers looked at the distractor picture (e.g., the picture of a *panda* upon hearing *pencil*) longer than pictures of names with distinct vowels (e.g., *strawberry* or *dice*). In contrast, L1 speakers did not look at the distractor picture more than the pictures of the other competitor words (e.g., *strawberry* or *dice*). Similar to the findings of Pallier and colleagues' study (2001), these findings also suggest that L2 participants' lower sensitivity to L2 phonological contrasts may interfere with the establishment of native-like phonological representations.

Although L2 learners' inability to discriminate L2 phonological contrasts that are not distinguished in the L1 poses a big challenge, that is not the only obstacle to the phonological encoding of L2 words. For L2 learners, learning the phonological forms of L2 words can still be challenging, even without difficult phonological contrasts. Meador, Flege and MacKay (2000) measured L1 and L2 speakers' probabilities of successful spoken word recognition in different levels of noise by asking participants to follow as many words as possible after the presentation of a sentence which always consisted of high-frequency words (i.e., words that should be well known to L2 speakers). The authors found that advanced Italian L2 speakers of English, despite a long residence in Canada (*M* = 35 years), were overall worse at word recognition than native speakers. In addition, although an L2 group with a low age-of-acquisition (AOA) (*M* = age 7) performed better than two other L2 groups with higher AOAs, this L2 group's performance was still significantly poorer than that of native speakers in the word recognition task, unless the signal-to-noise level was high. These discoveries indicate that L2 decoding ability is relatively impaired compared to L1 decoding ability. Therefore, it is possible to predict that the phonological forms of L2 words, which are usually acquired through exposure to natural aural input in noise (Gor et al., 2021; Luce & Pisoni, 1998), would be less likely to be accurately encoded and strengthened than L1 word forms.

Another important finding was made when Meador et al. (2000) examined the effect of early bilinguals' L1 use on their word recognition scores. The results showed that early bilinguals who seldom used Italian obtained significantly higher word recognition scores than early bilinguals who used their Italian relatively often. The authors assumed that good sound recognition requires accurate mental representations of phonetic segments. They then hypothesized that the detected L1 use effect might indicate the extent to which the L1 phonetic system affected the representations of L2 vowels and consonants. That is to say, they thought that the more the early bilinguals used L1, the more the representations for L1 phonetic segments influenced the representations that these early bilinguals had developed for L2 phonetic segments. This interpretation suggests that because adult L2 learners are not equipped with fine-grained representations of L2 phonetic segments due to their dominant use of the L1 during their childhood, accurate phonological encoding of L2 words that requires successful recognition of spoken words would be more challenging to them.

Unfaithful encoding of phonological forms of L2 words can also be affected by relatively poorer phonological short-term memory (PSTM) in the L2. N. Ellis (1994, 1996) suggested that learning the form of vocabulary involves sequencing the phonological properties of the language (the categorical units, syllable structure and phonotactic sequences). More specifically, he explained the role of PSTM in vocabulary acquisition: (a) repetition of sequences in PSTM allows their consolidation in phonological long-term memory (LTM), (b) the tuning of phonological LTM to regular sequences allows more ready perception of input that contains regular sequences, and (c) the cyclical reciprocal interactions of (a) and (b) allow learners to bootstrap their learning of L2 structure (Ellis, 1996, p. 108). Based on this reasoning, he predicted that individual differences in PSTM would predict learners' rate of word form acquisition, and this hypothesis turned out to be supported by empirical evidence (Martin & Ellis, 2012; Speciale, Ellis, & Bywater, 2004).

Considering the hypothesized roles of PSTM, it is reasonable to suppose that poor PSTM would inhibit phonological encoding of vocabulary. In Thorn and Gathercole's study (2001), the PSTM span was measured by an immediate recall task in which English-French bilingual participants were asked to speak out lists of words and nonwords that they had heard. The participants performed better in the recall task both when stimuli were words compared to when they were nonwords (both in the L1 and L2), and when stimuli were presented in the L1 compared to when they were presented in the L2. Based on their findings, the authors claimed that their participants' superior performance for word stimuli and L1 stimuli was because their use of long-term representations of phonological information, both of the phonological structure of specific lexical items and of the phonotactic properties of the language. Put differently, Thorn and Gathercole (2001) suspected that either long-term phonological representations would fill in the degraded information of the stimuli temporarily registered in short-term memory, or they would promote easier registration of familiar sound patterns in short-term memory. If the suggested mechanism of vocabulary acquisition that is supposed to be mediated by PSTM is recalled (Ellis, 1994, 1996), the relatively more limited PSTM span for L2 words compared to that for L1 words (Thorn & Gathercole, 2001) is predicted to be a reason for L2 learners' less successful encoding of the phonological forms of L2 words, even when these words are successfully perceived.

Meador and her colleagues' (Meador et al., 2000) and Thorn and Gathercole's (2001) studies suggest that L2 words may not be encoded as efficiently as L1 words, even when L2 words do not include confusing phonological contrasts, because of limited L2 decoding ability and poorer PSTM performance for L2 phonological signals. In consequence, it would be more difficult for L2 phonological representations to become fully specified than for L1 phonological representations if the same amount of input is provided. However, the FLR hypothesis (Cook & Gor, 2015; Gor, 2018, forthcoming; Gor & Cook, 2020; Gor et al., 2021) does not postulate that L2 phonological representations remain fuzzy forever. On the contrary, it posits that despite inefficient initial encoding, L2 representations can become robust enough to take part in lexical competition if correct L2 word forms are properly encoded through repeated exposure. In other words, the FLR hypothesis predicts that improved phonological encoding will first be observed for frequently encountered or familiar words that do not contain problematic L2 contrasts. This assumption is supported by Gor and Cook's (2020) finding that high-frequency prime-target pairs yielded form inhibition, unlike low-frequency pairs that produced form facilitation. Similarly, Cook and Gor (2015) showed that the primes which were perceived to be familiar by L2 participants produced form inhibition whereas other primes that were perceived to be recognizable but unfamiliar yielded form facilitation. These findings support the conjecture that the quality of phonological representation determines the pattern of form priming.

1.1.3. The Influence of Competitors

Before concluding that the patterns of form priming are determined by the quality of phonological representations in spoken word recognition, another important factor that influences form priming needs to be considered. Luce and Pisoni (1998), for instance, proposed that spoken word recognition is influenced by (1) the number and (2) degree of confusability of words in the neighborhood, as well as (3) the frequencies of the neighbors, when the neighbors were defined as the words with a deletion, insertion or substitution of one phoneme from a given target used in a spoken word LDT (for Experiments 2 and 3). Luce and Pisoni (1998) then demonstrated that the Neighborhood Activation Model (NAM), that they developed considering the effects of the three above-mentioned factors, was supported by the results of their experiments (Experiments 1 and 2). In the same vein, an effect of similarly sounding words (that are called competitors[5]) in a form-priming experiment was also discovered by Dufour and Peereman (2003a). These researchers showed that inhibitory form priming was stronger if the word target had a small competitor set size compared to when it had a large competitor set size. Dufour and Peereman (2003a) then interpreted their finding based on the following reasoning: (a) when the target had few competitors, the prime also had few competitors, and then (b) the prime with a small candidate size is more strongly activated because it is less influenced by its competitors as suggested by Luce and Pisoni (1998). (c) This prime then inhibits identification of the target more strongly.

If the number of phonological neighbors or competitors influences recognition of L1 words, it may also affect L2 word recognition. If so, it is possible to suppose that the competitor set size would be bigger for L2 words than for L1 words. This is because when hearing L2 words, even words that do not have very similar phonological forms are also activated, as shown by the activation of an L2 Catalan word /netə/ by another word /nɛtə/ which was observed in Spanish-dominant Spanish-Catalan bilinguals (Pallier et al., 2001). As noted earlier, this

---

[5] In spoken word identification research, the term *phonological neighbors* has been used only in restricted environments. For example, Luce and Pisoni (1998), who examined the effect of phonological neighbors, used only one-syllable words (e.g., *pat* or *cat*) in their study. The reason why this term has not been used frequently is that in spoken word recognition, a more important predictor of lexical competition is usually the word initial overlap between the phonological prime and the target rather than whether the prime is a phonological neighbor of the target (Maslen-Wilson, 1987). Competitors are a little different from phonological neighbors. "Competitor" refers to words that are supposed to compete with the target for selection such as the ones that overlap in the initial three phonemes with the target (e.g., *captive-captain*).

phenomenon was not observed if participants were Catalan-dominant bilinguals. Weber and Cutler (2004) also showed that when Dutch L2 learners of English heard the English word *panda*, a word containing a different vowel (/ɛ/ distinguished from /æ/), such as *pencil*, was also activated (in Experiment 1). Weber and Cutler (2004) further demonstrated that when L2 speakers heard the English word *kitten*, an L1 Dutch word *kist* (meaning "chest) was activated (in Experiment 3). In contrast, when L1 speakers heard the L1 word *kist*, a similarly sounding L2 word *kitten* was not activated (in Experiment 4). These findings seem to suggest that the number of competitors for L2 words is greater than that for L1 words because upon hearing L2 words, even less-similarly sounding L1 and L2 words are all spuriously activated and confused. However, Broersma (2012) showed that even though spurious activation of similarly sounding L2 words does occur, these activated L2 words do not take part in lexical competition. To be more specific, in a priming experiment (Experiment 1), Broersma (2012) used the nonword prime *daffo*- (which is the initial part of a word *daffodil*) and *defi*- (which was from *deficit*) for the word target *deficit*. Under this priming condition, it was expected that the influence of lexical competition would be controlled for because the nonword prime had been known to not compete for selection with the target. It was also predicted that *daffo*- would activate *deficit* only if the phoneme /æ/ is not distinguished from /ɛ/. Broersma (2012) then showed that both the primes *daffo*- and *defi*- produced facilitation for the target *deficit* in L2 speakers whereas only *defi*- yielded facilitation in L1 speakers. This observation indicates that spurious activation does occur in the L2. However, in Experiment 2 with a word prime that was predicted to trigger lexical competition, Broersma also discovered that the prime *flash* did not produce significant inhibition for the target *flesh* in the L2 but did in the L1. These results suggest that although L2 words may have a greater number of candidates, these candidates are not detrimental to word recognition because they do not strongly compete as L1 words do. Even though Broersma (2012) did not explicate the specific mechanism of these phenomena, her findings can be explained by the FLR hypothesis (Cook & Gor, 2015, Gor, 2018, forthcoming, Gor & Cook, 2020; Gor et al., forthcoming) which posits that words with weak lexical representations can be activated but do not work like words with strong lexical representations.

1.2. The Purpose of the Study

As introduced in detail in this chapter, the FLR hypothesis provides a theoretical framework that can explain the increased facilitation and reduced lexical competition observed in phonological-priming, spoken-word identification experiments (Cook & Gor, 2015; Gor & Cook, 2020). In addition, it was recently proposed that fuzziness may not reside only in the domain of phonological representations (Bordag et al, 2021; Gor, forthcoming; Gor et al, 2021), and several findings from previous visual word recognition research such as form confusion (e.g., misperception of *lamp* as *lamb*, Jiang & Zhang, 2021, p. 77) or form facilitation in priming experiments (see Section 2.1.2 below for a review of previous research) were newly interpreted as the results of supposedly less-specified L2 orthographic representations. However, greater support for the applicability of the FLR hypothesis to visual word processing may require further evidence. Therefore, with this purpose in mind, the current study aimed to test three key assumptions that were made based on the FLR hypothesis framework regarding written word recognition. First, we predicted that fuzzy orthographic representations would lead to decreased lexical competition between orthographic neighbors. Second, even though orthographic representations might become robust enough to take part in lexical competition at an earlier stage of lexical development, we expected that the fuzzy phonological representations of words with difficult phonological contrasts (e.g., *read-lead*) would still lead to distinctive processing of these words. Thus, words with nonnative contrasts were predicted to show weak lexical competition, even at a relatively later developmental stage (see Section 2.2.3 for the suspected reasons for this phenomenon). Third, we hypothesized that the quality of form (orthographic and phonological) representations would improve as a function of L2 experience. For that reason, nativelike performance was expected to eventually be observable in L2 speakers with precise orthographic and phonological representations. The next chapter reviews previous studies relevant to these assumptions.

Chapter 2. Literature Review

2.1. Orthographic Representations and Lexical Access

As the phonological priming method has been used to investigate the impact of phonological encoding to auditory lexical access, the orthographic priming method has also been used in visual word recognition studies to explore how the orthographic forms of words are encoded, or represented, in memory. In orthographic priming experiments, participants first encounter a prime that looks similar to the target. The researcher then examines how their response to the target differs from that after encountering an unrelated prime. Since the current study makes use of this orthographic priming method, it is necessary to understand what findings have been reported using this method in previous research and how these findings have been interpreted. Information about under which condition a certain pattern of priming has been observed is also needed in order to choose the proper stimuli and priming procedure and not to misinterpret results that are obtained in the current study. Therefore, this chapter reviews the literature with these purposes in mind.

2.1.1. Form Priming Experiments in the L1

In orthographic priming experiments, the PLE has been observed just as in phonological priming research (see Section 1.1.1). For example, Davis and Lupker (2006) conducted a series of orthographic neighbor priming experiments using the masked priming paradigm, where a neighbor is defined as a word that differs from the word target by one letter (e.g., *axle-able*). In this paradigm, the prime is preceded by a forward mask (####) and is immediately followed by the target stimulus, which is classified as either a word or a nonword. Under this masked condition, and with a very short prime-target stimulus onset asynchrony (SOA) of 30-60 ms, the prime is usually not consciously identifiable, and yet it has marked effects on the processing of the target word (Forster & Davis, 1984). Davis and Lupker (2006) discovered that when a word target was preceded by a masked orthographic neighbor (e.g., *axle-ABLE*), native speakers' response to the target in an LDT was delayed compared to when it was preceded by an unrelated word prime (e.g., *thug-ABLE*). This contrasted with the same participants' faster reaction time to a target that was preceded by a form-related nonword prime (e.g., *ible-ABLE*) compared to when it preceded by an unrelated nonword prime (e.g., *shug-ABLE*).

The PLE found in written word recognition studies can be explained by interaction-based models such as those in phonological priming spoken-word recognition research. For instance, according to the Interactive Activation (IA) model suggested by McClelland and Rumelhart (1981), the PLE is understood as the outcome of lexical competition between the word prime and the word target, and of the absence of lexical competition between the nonword prime and the word target (e.g., Colombo, 1986; Davis & Lupker, 2006). To be more specific, the IA model posits that there are three levels of representation, a feature level, a letter level, and a word level. These levels are connected through facilitatory and inhibitory pathways between levels, and there are additional inhibitory connections within the letter level and within the word level. For example, when a participant sees the word *axle*, a vertical line "|" activates all letters that have this feature including "a" while inhibiting the activation of other letters that do not have this feature such as "s". After that, "a" activates all words having this letter in the first letter position, and once many words are activated simultaneously, intra-level inhibition then begins to emerge to inhibit the activation of all other words that start with "a" except for *axle*. Therefore, recognition of its orthographic neighbor *able* is inhibited. However, there is no intra-level inhibition between a nonword and a word, but there is inter-level facilitation, so recognition of the word target is boosted when it is primed by a form-related nonword such as *ible*.

The PLE can also be explained by the so-called *best-match hypothesis* that was proposed based on the entry opening model (Forster, 1987; Forster & Davis, 1984; Forster, Davis, Schoknecht, & Carter, 1987; Forster & Veres, 1998; Qiao & Forster, 2013). This account posits that lexical entries should be "opened" before any information can be retrieved from them, and this process is hypothesized to require some time. If the resolution process for the prime (i.e., identification of the prime) is not complete before the presentation of the target, this facilitates lexical access to the target since the prime opens all lexical entries that have similar forms. In other words, the word identification process for the target gets a head start at the moment when the prime is encountered. However, sometimes the resolution process for the prime is finished before the target is encountered. In that case, the prime does not produce facilitation, and occasionally produces inhibition because a closed-down lexical entry remains in a refractory state for a while. Based on this mechanism, the nonword prime is understood to produce facilitation since the resolution process for the prime cannot be completed quickly, given that no lexical entry matches it. On the other hand, because identification of the word prime is quickly achieved, it is thought to not facilitate target recognition.

These two accounts may sound straightforward, but research suggests that the mechanism for form priming could be far more complex than it was assumed to be. The following section will briefly review the findings of orthographic neighbor priming studies in the L1.

The Word Target in Unmasked Priming Studies. As shown in Table 1, only one study (Forster & Veres, 1998) has investigated whether the PLE is observed using both orthographic neighbor primes and form-related nonword primes in a single study, and it confirmed the existence of this effect. A comparison of several studies (Forster & Veres, 1998; Park, 2021; Segui & Grainger, 1990) suggests that inhibition is stronger when the length of a stimulus is shorter relative to when it is longer. The weak inhibition for long words can be thought to have come from strong inter-level excitatory activation caused by many-letter overlap between the prime and target (Elgort, 2011; Forster & Veres, 1998). However, weak inhibition for long words could also come from the effect of low neighborhood density because it is known that inhibition is weak if words have few neighbors (Forster et al., 1987; Nakayama, Sears, & Lupker, 2008, and see below), and long words tend to have fewer neighbors than short words do.

Another discovery from the previous research is that when the prime was visible, inhibition was stronger if prime frequency was low compared to if it was high (Colombo, 1986; Segui & Grainger, 1990). This phenomenon is suspected to occur because high-frequency targets must be inhibited to a greater extent in order for low-frequency primes to be successfully identified when the prime is encountered (i.e., before the target is presented). Colombo (1986) also compared the strength of priming from primes that shared the last 3-5 letters with the target to that from other primes that shared the initial 2-3 letters (in Experiment 2 and 3). She then found that in the latter case, inhibition occurred regardless of target frequency whereas in the former case, only high-frequency targets showed inhibition. Colombo (1986) interpreted that this is because when the initial letters overlap, relevant words (i.e., words that have similar forms to the prime including the target) are activated faster regardless of their resting levels, whereas when the rhymes (i.e., last letters) overlap, words are slowly activated unless they are high-frequency words. The last two findings (strong inhibition for low-frequency prime–high-frequency target pairs and for prime-target pairs with initial overlap) are consistent with arguments proposed in spoken word recognition research concerning the phonological priming effect (Eberhard, 1994; Marslen-Wilson, 1987).

The Nonword Target in Unmasked Priming Studies. Forster and Veres (1998) report that word primes produced significant facilitation for nonword targets, although nonword primes did not. However, in a study that used the medium-term priming method (i.e., a study in which participants make lexical decisions both on the prime and the target that are separated by other lexical decision stimuli), Park (2021) did not find the same priming effect. These contrasting findings suggest that the facilitative priming originating from awareness of a formal similarity between the word prime and the nonword target may disappear very shortly, and the priming that endures a substantial time gap occurs only when the prime pre-activates the lexical representation of the target word.

To sum up, unmasked orthographic priming studies have shown findings that are surprisingly similar to those of phonological priming studies. Thus, these findings of unmasked orthographic priming experiments support the assumption that form-based facilitation is a general modality-independent property of word recognition (Gor & Cook, 2020). However, masked orthographic priming studies have shown less straightforward and more complicated results. The following section will review the findings of this masked priming research.

Table 1

The Results of Unmasked Orthographic Form Priming Experiments in the L1

Study	**Prime lexicality**	**Prime duration**	**Length**	**Target lexicality**	**Target duration**	**Target N density**	**RT: Related**	**RT: Unrelated**	**Unrelated - related**	**Special manipulations**	**Language**
Colombo (1986) Experiment 1	Word	240	4~7	Word	Until being responded	Not reported	600	602	2	The prime and the target shared rhymes (3-5 letters), No stats.	Italian
Colombo (1986) Experiment 1	Nonword	240	4~7	Word	Until being responded	Not reported	629	606	-23		Italian
Colombo (1986) Experiment 1	Word	240	4~7	Nonword	Until being responded	Not reported	684	678	-6		Italian
Colombo (1986) Experiment 1	Nonword	240	4~7	Nonword	Until being responded	Not reported	683	706	23		Italian
Colombo (1986) Experiment 1	Word	640	4~7	Word	Until being responded	Not reported	641	620	-21		Italian
Colombo (1986) Experiment 1	Nonword	640	4~7	Word	Until being responded	Not reported	625	610	-15		Italian
Colombo (1986) Experiment 1	Word	640	4~7	Nonword	Until being responded	Not reported	690	713	23		Italian
Colombo (1986) Experiment 1	Nonword	640	4~7	Nonword	Until being responded	Not reported	726	740	14		Italian
Colombo (1986) Experiment 2	Word	320	4~7	Word (High F)	Until being responded	Not reported	599	548	-51*	The prime and the target shared rhymes (3-5 letters).	Italian
Colombo (1986) Experiment 2	Word	320	4~7	Word (Low F)	Until being responded	Not reported	644	699	55*		Italian
Colombo (1986) Experiment 3	Word	320	4~7	Word (High F)	Until being responded	Not reported	601	557	-44*	The prime and the target shared the initial letters (2-3 letters).	Italian
Colombo (1986) Experiment 3	Word	320	4~7	Word (High F)	Until being responded	Not reported	603	557	-46*		Italian
Colombo (1986) Experiment 3	Word	320	4~7	Word (Low F)	Until being responded	Not reported	714	690	-24		Italian
Colombo (1986) Experiment 3	Word	320	4~7	Word (Low F)	Until being responded	Not reported	710	690	-20		Italian
Segui & Grainger (1990) Experiment 1	Word (Low F)	350	4	Word (High F)	Until being responded	Not reported	630	598	-32*	Neighborhood frequency effect	French
Segui & Grainger (1990) Experiment 1	Word (High F)	350	4	Word (Low F)	Until being responded	Not reported	687	702	15		French
Segui & Grainger (1990) Experiment 3	Word (Low F)	350	4	Word	Until being responded	Not reported	639	605	-34*	Neighborhood frequency effect	Dutch
Segui & Grainger (1990) Experiment 3	Word (High F)	350	4	Word	Until being responded	Not reported	609	611	2		Dutch
Forster & Veres (1998) Experiment 1	Word	500	8~9	Word	500	1.21	850	853	3	Nonwords were one-letter different from real words.	English
Forster & Veres (1998) Experiment 1	Nonword	500	8~9	Word	500	1.04/1.21	795	853	58*		English
Forster & Veres (1998) Experiment 1	Word	500	8~9	Nonword	500	1.04/1.21	940	1006	66*		English
Forster & Veres (1998) Experiment 1	Nonword	500	8~9	Nonword	500	1.04	1040	1006	-34		English
Park (2021)	Word (Low F)	Until being responded	6~10	Word	Until being responded	1.7	630	633	3	The frequency of the prime was higher than that of the target.	English
Park (2021)	Word (High F)	Until being responded	6~10	Word	Until being responded	1.7	631	621	-10		English
Park (2021)	Word	Until being responded	6~10	Nonword	Until being responded	1	666	664	-2		English

F: Frequency, N: neighbor, No stats.: Reporting main effects only without statistical analysis for the form priming effect under each condition, *: $p < .05$.

Table 2

The Results of Masked Orthographic Form Priming Experiments in the L1

Study	Prime lexicality	Prime duration	Length	Target lexicality	Target duration	Target N density	RT: Related	RT: Unrelated	Unrelated - related	Special manipulations	Language
Forster & Davis (1984) Experiment 2	Word	50	4	Word	500	Not reported	524	523	-1		English
Forster & Davis (1984) Experiment 2	Nonword	50	5	Nonword	500	Not reported	552	559	7		English
Forster et al. (1987) Experiment 1	Nonword (High F, transposed)	60	6~9	Word	500	Not reported	464	527	63*		English
Forster et al. (1987) Experiment 1	Nonword (High F, substituted)	60	6~9	Word	500	Not reported	477	527	50*		English
Forster et al. (1987) Experiment 1	Nonword (Low F, transposed)	60	6~9	Word	500	Not reported	510	572	62*		English
Forster et al. (1987) Experiment 1	Nonword (Low F, substituted)	60	6~9	Word	500	Not reported	523	572	49*		English
Forster et al. (1987) Experiment 1	Nonword	60	6~9	Nonword	500	Not reported	566	570	4		English
Forster et al. (1987) Experiment 1	Nonword	60	6~9	Nonword	500	Not reported	559	570	11		English
Forster et al. (1987) Experiment 2	Nonword	60	4	Word	500	Not reported	492	500	8	Long words produce facilitation.	English
Forster et al. (1987) Experiment 2	Nonword	60	8	Word	500	Not reported	490	522	32*		English
Forster et al. (1987) Experiment 2	Nonword	60	4	Nonword	500	Not reported	547	547	0		English
Forster et al. (1987) Experiment 2	Nonword	60	8	Nonword	500	Not reported	558	557	-1		English
Forster et al. (1987) Experiment 4	Nonword	60	4	Word (High F)	500	Not reported	499	492	-7		English
Forster et al. (1987) Experiment 4	Nonword	60	4	Word (Low F)	500	Not reported	531	543	12		English
Forster et al. (1987) Experiment 4	Nonword	60	4	Nonword	500	Not reported	558	570	12		English
Forster et al. (1987) Experiment 5	Nonword	60	4	Word (Low N)	500	2.86	556	578	22*		English
Forster et al. (1987) Experiment 5	Nonword	60	4	Nonword	500	2.86	597	606	9		English
Forster et al. (1987) Experiment 6	Nonword	60	4	Word (Low N)	500	12.6	467	451	-16	A high N density produces stronger inhibition.	English
Forster et al. (1987) Experiment 6	Nonword	60	4	Word (High N)	500	1.6	459	477	18*		English
Forster et al. (1987) Experiment 7	Word (Morphologically related)	60	4	Word	500	7.96	491	527	36*	e.g., *keep-kept*	English
Forster (1987) Experiment 1	Word	60	8 <	Word	500	Not reported	458	496	38*		English
Forster (1987) Experiment 1	Nonword	60	8 <	Word	500	Not reported	468	496	28*		English
Forster (1987) Experiment 2	Nonword (With few Ns)	60	4	Word	500	9.14	482	489	7		English
Forster (1987) Experiment 2	Nonword (With high N)	60	4	Word	500	9.14	480	489	9		English
Forster (1987) Experiment 3	Nonword (With a backward mask)	60	8	Word	500	Not reported	525	544	19*		English

Study	Prime lexicality	Prime duration	Length	Target lexicality	Target duration	Target N density	RT: Related	RT: Unrelated	Unrelated - related	Special manipulations	Language
Segui & Grainger (1990) Experiment 2	Word (Low F)	60	4	Word (High F)	Until being responded	Not reported	598	608	10	N frequency effect	French
Segui & Grainger (1990) Experiment 2	Word (High F)	60	4	Word (Low F)	Until being responded	Not reported	709	661	-48*		French
Segui & Grainger (1990) Experiment 3	Word (Low F)	60	4	Word	Until being responded	Not reported	643	631	-12	N frequency effect	Dutch
Segui & Grainger (1990) Experiment 3	Word (High F)	60	4	Word	Until being responded	Not reported	662	621	-41*		Dutch
Ferrand & Grainger (1992) Experiment 1	Nonword	64	4	Word (High F)	Until being responded	Not reported	601	605	4	No stats.	French
Ferrand & Grainger (1992) Experiment 1	Nonword	64	4	Word (Low F)	Until being responded	Not reported	641	644	3	No stats.	French
Ferrand & Grainger (1992) Experiment 1	Nonword	64	4	Nonword	Until being responded	Not reported	692	697	5	No stats.	French
Ferrand & Grainger (1992) Experiment 2	Nonword	64	4	Word (High F)	Until being responded	Not reported	608	603	-5	Use of pseudohomophone nonwords, No stats.	French
Ferrand & Grainger (1992) Experiment 2	Nonword	64	4	Word (Low F)	Until being responded	Not reported	640	638	-2		French
Ferrand & Grainger (1992) Experiment 2	Nonword	64	4	Nonword (Pseudohomophone)	Until being responded	Not reported	738	750	12		French
Ferrand & Grainger (1992) Experiment 2	Nonword	64	4	Nonword	Until being responded	Not reported	680	704	24		French
Ferrand & Grainger (1992) Experiment 3	Nonword	32	4	Word (High F)	Until being responded	Not reported	563	593	30	No stats. Significant inhibition (a main effect)	French
Ferrand & Grainger (1992) Experiment 3	Nonword	32	4	Word (Low F)	Until being responded	Not reported	604	620	16		French
Ferrand & Grainger (1992) Experiment 3	Nonword	32	4	Nonword	Until being responded	Not reported	633	638	5		French
Grainger & Ferrand (1994) Experiment 2	Word	64	4~5	Word	Until being responded	Not reported	671	640	-31	Without psudohomophone nonwords	French
Grainger & Ferrand (1994) Experiment 2	Word	64	4~5	Word	Until being responded	Not reported	748	717	-31	With psudohomophone nonwords	French
Grainger & Ferrand (1994) Experiment 3	Word	64	3~6	Word	Until being responded	Not reported	670	647	-23		English
Bijeljac-Babic et al. (1997) Experiment 2	Word	57	4	Word	Until being responded	Not reported	740	690	-50	Monolingual participants, no stats.	French
Bijeljac-Babic et al. (1997) Experiment 2	Word	57	4	Word	Until being responded	Not reported	785	734	-51	Low-prof. bilingual participants, no stats.	French
Bijeljac-Babic et al. (1997) Experiment 2	Word	57	4	Word	Until being responded	Not reported	757	729	-28	High-prof. bilingual participants, no stats.	French

Study	Prime lexicality	Prime duration	Length	Target lexicality	Target duration	Target N density	RT: Related	RT: Unrelated	Unrelated - related	Special manipulations	Language
Forster & Veres (1998) Experiment 2	Word	50	8~9	Word	500	1.21	701	709	8	Nonwords were one-letter different from real words	English
Forster & Veres (1998) Experiment 2	Nonword	50	8~9	Word	500	1.04/1.21	672	709	37*		English
Forster & Veres (1998) Experiment 2	Word	50	8~9	Nonword	500	1.04/1.21	834	852	18		English
Forster & Veres (1998) Experiment 2	Nonword	50	8~9	Nonword	500	1.04	826	852	26		English
Forster & Veres (1998) Experiment 3	Word	50	8~9	Word	500	1.21	540	574	34*	Nonwords did not look like words.	English
Forster & Veres (1998) Experiment 3	Nonword	50	8~9	Word	500	1.04/1.21	549	574	25*		English
Forster & Veres (1998) Experiment 4	Word	50	8~9	Word	500	1.21	623	655	32*	Nonwords were two-letter different from real words	English
Forster & Veres (1998) Experiment 4	Nonword	50	8~9	Word	500	1.04/1.21	621	655	34*		English
Forster & Veres (1998) Experiment 4	Word	50	8~9	Nonword	500	1.04/1.21	745	744	-1		English
Forster & Veres (1998) Experiment 4	Nonword	50	8~9	Nonword	500	1.04	755	744	-11		English
Castle et al. (1999)	Nonword	57	4~5	Word (High N)	800	7.9	601	601	0	Adult participants	English
Castle et al. (1999)	Nonword	57	4~5	Word (Low N)	800	1.3	580	593	13		English
De Moor & Brysbaert (2000)	Word (Same length)	57	4~5	Word	Until being responded	Not reported	664	650	-14	The effect of the orthographic N prime with different lengths	Dutch
De Moor & Brysbaert (2000)	Word (Different length)	57	4~5	Word	Until being responded	Not reported	708	687	-21*		Dutch
De Moor & Brysbaert (2000)	Word (Same length)	57	4~5	Nonword	Until being responded	Not reported	738	737	-1		Dutch
De Moor & Brysbaert (2000)	Word (Different length)	57	4~5	Nonword	Until being responded	Not reported	754	770	16		Dutch
De Moor et al. (2005) Experiment 1	Word	57	4~5	Word (Offline feedback)	Until being responded	Not reported	675	662	-13	Stressing accuracy	Dutch
De Moor et al. (2005) Experiment 1	Word	57	4~5	Word (Online feedback)	Until being responded	Not reported	686	641	-45*		Dutch
De Moor et al. (2005) Experiment 1	Word	57	4~5	Nonword (Offline feedback)	Until being responded	Not reported	705	724	19		Dutch
De Moor et al. (2005) Experiment 1	Word	57	4~5	Nonword (Online feedback)	Until being responded	Not reported	722	724	2		Dutch
De Moor et al. (2005) Experiment 2	Word	57	4~5	Word (Offline feedback)	Until being responded	Not reported	637	622	-15	Stressing speed	Dutch
De Moor et al. (2005) Experiment 2	Word	57	4~5	Word (Online feedback)	Until being responded	Not reported	498	526	28*		Dutch
De Moor et al. (2005) Experiment 2	Word	57	4~5	Nonword (Offline feedback)	Until being responded	Not reported	675	665	-10		Dutch
De Moor et al. (2005) Experiment 2	Word	57	4~5	Nonword (Online feedback)	Until being responded	Not reported	539	520	-19		Dutch

Study	Prime lexicality	Prime duration	Length	Target lexicality	Target duration	Target N density	RT: Related	RT: Unrelated	Unrelated - related	Special manipulations	Language
Davis & Lupker (2006) Experiment 1	Word	57	4~5	Word (High F)	Until being responded	2.2	679	645	-34	No stats. Significant inhibition from word primes (a main effect)	English
Davis & Lupker (2006) Experiment 1	Nonword	57	4~5	Word (High F)	Until being responded	2.2	634	660	26		English
Davis & Lupker (2006) Experiment 1	Word	57	4~5	Word (Low F)	Until being responded	2.2	586	573	-13		English
Davis & Lupker (2006) Experiment 1	Nonword	57	4~5	Word (Low F)	Until being responded	2.2	571	582	11		English
Davis & Lupker (2006) Experiment 1	Word	57	4~5	Nonword	Until being responded	2.2	743	757	14		English
Davis & Lupker (2006) Experiment 1	Nonword	57	4~5	Nonword	Until being responded	2.2	737	744	7		English
Davis & Lupker (2006) Experiment 2	Word	57	5	Word	Until being responded	3.5	736	729	-7	The effect of shared Ns between the prime and the target	English
Davis & Lupker (2006) Experiment 2	Word (Without shared neighbors)	57	5	Word	Until being responded	3.5	720	685	-35*		English
Davis & Lupker (2006) Experiment 2	Word (With a shared neighbor)	57	5	Nonword	Until being responded	3.5	826	826	0		English
Davis & Lupker (2006) Experiment 3	Word	57	4	Word (High N)	Until being responded	13.1	659	638	-21	Word target with high-N nonwords (N = 13.2).No stats.	English
Davis & Lupker (2006) Experiment 3	Word	57	4	Word (Low N)	Until being responded	2.8	684	647	-37		English
Davis & Lupker (2006) Experiment 3	Word	57	4	Word (High N)	Until being responded	13.1	615	607	-8	Word target with Low-N nonwords (N = 2.8). No stats.	English
Davis & Lupker (2006) Experiment 3	Word	57	4	Word (Low N)	Until being responded	2.8	649	632	-17		English
Davis & Lupker (2006) Experiment 3	Word	57	4	Nonword (High N)	Until being responded	13.2	784	786	2	No stats.	English
Davis & Lupker (2006) Experiment 3	Word	57	4	Nonword (Low N)	Until being responded	2.8	737	713	-24		English
Castle et al. (2007)	Nonword (Substituted)	57	4~5	Word	800	6.1	576	583	7	Adult participants	English
Castle et al. (2007)	Nonword (Transposed)	57	4~5	Word	800	6.1	575	583	8		English
De Moor et al. (2007)	Word	14	4	Word	Until being responded	Not reported	680	689	9	The effect of prime duration, no stats. A significant inhibition for word prime and word target pairs (a main effect)	Dutch
De Moor et al. (2007)	Word	29	4	Word	Until being responded	Not reported	690	671	-19		Dutch
De Moor et al. (2007)	Word	43	4	Word	Until being responded	Not reported	780	761	-19		Dutch
De Moor et al. (2007)	Word	57	4	Word	Until being responded	Not reported	815	750	-65		Dutch
De Moor et al. (2007)	Word	14	4	Nonword	Until being responded	Not reported	694	685	-9		Dutch
De Moor et al. (2007)	Word	29	4	Nonword	Until being responded	Not reported	703	690	-13		Dutch
De Moor et al. (2007)	Word	43	4	Nonword	Until being responded	Not reported	745	764	19		Dutch
De Moore et al. (2007)	Word	57	4	Nonword	Until being responded	Not reported	741	764	23		Dutch

Study	Prime lexicality	Prime duration	Length	Target lexicality	Target duration	Target N density	RT: Related	RT: Unrelated	Unrelated - related	Special manipulations	Language
Nakayama et al. (2008) Experiment 1	Word (High-freq.)	60	4	Word (Low F)	Until being responded	9.8	618	594	-24	Relative F of the prime and the target with many neighbors. Significant inhibition (a main effect).	English
Nakayama et al. (2008) Experiment 1	Word (Low-freq.)	60	4	Word (High F)	Until being responded	9.8	537	516	-21		English
Nakayama et al. (2008) Experiment 2	Word (High-freq.)	60	4~5	Word (Low F)	Until being responded	10.3	613	573	-40	Relative F of the prime and the target with many neighbors. Significant inhibition (a main effect).	English
Nakayama et al. (2008) Experiment 2	Word (Low-freq.)	60	4~5	Word (High F)	Until being responded	10.1	548	523	-25		English
Nakayama et al. (2008) Experiment 2	Word (High-freq.)	60	4~5	Nonword	Until being responded	10.2	634	634	0		English
Nakayama et al. (2008) Experiment 2	Word (Low-freq.)	60	4~5	Nonword	Until being responded	10.1	623	632	9		English
Nakayama et al. (2008) Experiment 3	Word (High-freq.)	60	4~5	Word (Low F)	Until being responded	9.6	585	558	-27*	Word targets with many Ns	English
Nakayama et al. (2008) Experiment 3	Word (Low-freq.)	60	4~5	Word (High F)	Until being responded	9.2	549	517	-32*		English
Nakayama et al. (2008) Experiment 3	Word (High-freq.)	60	4~5	Word (Low F)	Until being responded	2.9	598	580	-18*	Word targets with few Ns	English
Nakayama et al. (2008) Experiment 3	Word (Low-freq.)	60	4~5	Word (High F)	Until being responded	2.4	529	529	0		English
Nakayama et al. (2008) Experiment 3	Word (High-freq.)	60	4~5	Nonword	Until being responded	8.5	640	631	9	Nonword targets with many Ns	English
Nakayama et al. (2008) Experiment 3	Word (Low-freq.)	60	4~5	Nonword	Until being responded	9.7	648	659	11		English
Nakayama et al. (2008) Experiment 3	Word (High-freq.)	60	4~5	Nonword	Until being responded	2.6	614	619	5	Nonword targets with few Ns	English
Nakayama et al. (2008) Experiment 3	Word (Low-freq.)	60	4~5	Nonword	Until being responded	2.8	622	625	-3		English
Nakayama et al. (2008) Experiment 4A	Word (With a shared N)	60	4~5	Word	Until being responded	11.1	568	547	-21*	Targets with many Ns	English
Nakayama et al. (2008) Experiment 4A	Word (Without shared Ns)	60	4~5	Word	Until being responded	11.1	568	547	-21*		English
Nakayama et al. (2008) Experiment 4A	Word (Without shared Ns)	60	4~5	Nonword	Until being responded	9.4	664	668	4		English
Nakayama et al. (2008) Experiment 4B	Word (With a shared N)	61	4~5	Word	Until being responded	3.5	575	572	-3	Targets with few Ns	English
Nakayama et al. (2008) Experiment 4B	Word (Without shared Ns)	62	4~5	Word	Until being responded	2.6	556	550	-6		English
Nakayama et al. (2008) Experiment 4B	Word (Many Ns)	63	4~5	Nonword	Until being responded	10	682	687	5		English
Nakayama et al. (2008) Experiment 4B	Word (Few Ns)	64	4~5	Nonword	Until being responded	3.1	663	645	-18		English

Study	Prime lexicality	Prime duration	Length	Target lexicality	Target duration	Target N density	RT: Related	RT: Unrelated	Unrelated - related	Special manipulations	Language
Andrews & Hersch (2010) Experiment 1	Word	50	4	Word (High N)	500	12.5	609	613	4	No stats.	English
Andrews & Hersch (2010) Experiment 1	Nonword	50	4	Word (High N)	500	12.5	602	603	1		English
Andrews & Hersch (2010) Experiment 1	Word	50	4	Word (Low N)	500	3.7	625	629	4		English
Andrews & Hersch (2010) Experiment 1	Nonword	50	4	Word (Low N)	500	3.7	594	612	18		English
Andrews & Hersch (2010) Experiment 1	Word	50	4	Nonword (High N)	500	Not reported	720	714	-6		English
Andrews & Hersch (2010) Experiment 1	Word	50	4	Nonword (Low N)	500	Not reported	696	697	1		English
Andrews & Hersch (2010) Experiment 2	Word	50	5	Word (High N)	500	6.7	650	628	-22	No stats.	English
Andrews & Hersch (2010) Experiment 2	Nonword (Ambiguous partial)	50	5	Word (High N)	500	6.7	620	628	8		English
Andrews & Hersch (2010) Experiment 2	Nonword (Unambiguous partial)	50	5	Word (High N)	500	6.7	614	628	14		English
Andrews & Hersch (2010) Experiment 2	Word	50	5	Word (Low N)	500	1.6	623	625	2		English
Andrews & Hersch (2010) Experiment 2	Nonword (Ambiguous partial)	50	5	Word (Low N)	500	1.6	606	625	19		English
Andrews & Hersch (2010) Experiment 2	Nonword (Unambiguous partial)	50	5	Word (Low N)	500	1.6	604	625	21		English
Andrews & Hersch (2010) Experiment 2	Word	50	5	Nonword (High N)	500	Not reported	722	729	7		English
Andrews & Hersch (2010) Experiment 2	Word	50	5	Nonword (Low N)	500	Not reported	776	785	9		English
Nakayama & Lupker (2018) Experiment 2	Word	67	4~5	Word	Until being responded	5.5	582	560	-22*		English
Nakayama & Lupker (2018) Experiment 2	Nonword	67	4~5	Word	Until being responded	5.5	568	558	-10		English
Nakayama & Lupker (2018) Experiment 2	Word	67	4~5	Nonword	Until being responded	5.1	664	630	-34*		English
Nakayama & Lupker (2018) Experiment 2	Nonword	67	4~5	Nonword	Until being responded	5.1	627	640	13		English
Jiang (2021)	Word	50	5~7	Word	Until being responded	2.3	549	560	11		English
Park (2021)	Word (Low F)	67	4~5	Word	Until being responded	9.2	699	663	-36*		English
Park (2021)	Word (High F)	67	4~5	Word	Until being responded	9.6	689	651	-38*		English
Park (2021)	Word (Low F)	67	4~5	Nonword	Until being responded	6.6	754	753	-1		English
Park (2021)	Word (High F)	67	4~5	Nonword	Until being responded	6.3	732	747	15		English

F: Frequency, N: neighbor, No stats.: Reporting main effects only without statistical analysis for the form priming effect under each condition, *: *p* < .05.

The Word Target in Masked Priming Studies. As shown in Table 2, a greater number of studies have been conducted using the masked priming paradigm than the unmasked priming paradigm since it is known that this method is less influenced by strategic effects (i.e., the effects of participants' awareness of the relationship between the prime and the target) and can be used to explore unconscious and automatic cognitive processes taking place during visual word recognition. The studies illustrated in Table 2 show that the patterns of priming are determined by an interplay of many different factors. Inhibition from the prime was stronger if shorter stimuli were used (Forster, 1987; Forster et al., 1987), if the prime or target had many neighbors (Andrews & Hersch, 2010; Davis & Lupker, 2006; Forster et al., 1987; Nakayama, Sears, & Lupker, 2008), if the prime and target shared neighbors (Davis & Lupker, 2006; but also see Nakayama et al., 2008) and if the frequency of the prime was higher (Nakayama et al., 2008; Segui & Grainger, 1990). It is also found that the strength of form inhibition is affected by prime duration (Davis & Lupker, 2006; De Moor, Van der Herten, & Verguts, 2007; Ferrand & Grainger, 1992), the levels of difficulty in discrimination of word targets from nonword foils (De Moor, Van der Herten, & Verguts, 2007; Forster & Veres, 1998) and whether accuracy is stressed over speed or vice versa (De Moor, Verguts, & Brysbaert, 2005).

In order to explain how the above-mentioned factors influence form priming, attempts to refine initial word identification models have been made. For instance, within the framework of interaction-based models, it has been suggested that the greater inhibition from high-frequency primes is because of the low resting levels of relatively lower-frequency targets that make them more susceptible to interference from strong competitors (i.e., a high-frequency prime or other high-frequency neighbors). To be more specific, it is assumed that when the masked word prime is encountered, only high-frequency neighbors (including the target itself) are more strongly activated than low-frequency words. Thus, these strongly activated high-frequency primes are thought to exert greater influence on their weak competitors (low-frequency neighbors including the target), while weakly activated low-frequency primes do not fiercely compete with other candidates for selection (Nakayama et al., 2008; Segui & Grainger, 1990). In the same vein, the tendency for word stimuli with many neighbors to produce stronger inhibition is interpreted as the outcome of the coactivation of neighbors in a high-density neighborhood, collectively producing large inhibition and/or the existence of a strong competitor for the target (i.e., a high-frequency prime or another high-frequency neighbor) (Nakayama et al., 2008).

In the case of the best-match hypothesis, the *lexical tuning hypothesis* was proposed in order to accommodate the effect of the prime/target neighborhood density in masked priming studies, (Castles, Davis, & Letcher, 1999; Castles et al., 2007; Forster & Veres, 1998; Qiao & Forster, 2013). This hypothesis posits that as more and more words with similar forms are represented in the mental lexicon along with the growth of one's vocabulary, these words are gradually better tuned in order not to be confused with their neighbors. For that reason, orthographic representations of words that have many neighbors are more narrowly tuned compared to those of words with few neighborhoods. Therefore, when word stimuli with few neighbors are used in masked priming experiments, facilitation is often observed. This facilitation supposedly takes place because broadly tuned orthographic representations are easily activated, even when the forms of the prime and the target are not exactly matched. Another proposed reason for form facilitation for word prime-target pairs with few neighbors is the unsuccessful resolution of the prime within a short prime duration due to its less precise orthographic representation. This interpretation is based on the understanding that the quality of orthographic representations determines the speed of resolution of the masked prime (Andrews

& Hersch, 2010; Andrews & Lo, 2012; Perfetti & Hart; 2002). Unsuccessful verification of primes with few neighbors, therefore, is believed to open the entry of the target, which facilitates target recognition. Meanwhile, the null effect from word primes with many neighbors is considered the result of either non-activation of the target by its not-exactly-matched neighbor prime, or quick verification of the prime that swiftly closes the entries of its neighbors. At the same time, form inhibition is thought to occur because once checked and rejected, candidates during the prime resolution process are more difficult to retrieve in subsequent target identification.

Another important concept in form priming research, the PLE, has been observed in several studies (Andrews & Hersch, 2010; Forster & Veres, 1998; Davis & Lupker, 2006). Table 2 shows that the strength of inhibition from a nonword prime overall tends to be weaker than that from a word prime, such that significant inhibition from the nonword prime has never been detected in any of the studies, whereas significant facilitation from the word prime has never been observed except in one study (Forster & Veres, 1998) in which long words with few neighbors were used as stimuli. However, a close inspection of the previous studies provides a more important implication: the realization that form priming may need to be understood as a continuum rather than the product of an on-off switch of prime lexicality. That is to say, in a form priming experiment, null priming for word targets does not necessarily indicate the complete absence of lexical competition because it could still be the outcome of complicated interactions between the prime, the target and their orthographic neighbors (see Table 2). Thus, if a null priming effect from a form prime is observed in the present study, according to the interaction-based models, it should be regarded as the result of weak lexical competition, as in previous studies (Castles et al., 2007; Elgort, 2011; Elgort & Warren, 2014; Jiang, 2021).

Alternatively, based on the lexical tuning hypothesis, a null priming effect could be understood as the outcome of precise orthographic representations, which does not *always* lead to entry opening of the target (i.e., form facilitation).

The Nonword Target in Masked Priming Studies. The priming effect for the nonword target has often shown great variability, so even a numerically large priming effect was sometimes non-significant (e.g., 26 ms of facilitation found in Experiment 2 of the Forster and Veres study was not significant). Thus, neither the word nor the nonword primes produced significant priming except in only one study (Nakayama & Lupker, 2018, Experiment 2). Since both interaction-based models and the lexical tuning (or the best-match) hypothesis make no prediction for nonword targets, it is difficult to interpret the significant inhibitory priming observed in Nakayama and Lupker's study (2018), as the authors also admitted.

Development of Orthographic Representations and Its Impact on Form Priming. The lexical tuning hypothesis suggests a developmental nature for lexical representations (Castle et al., 1999, 2007). Specifically speaking, it posits that early in reading development, when children's sight vocabularies are small, orthographic representations of words are as broadly tuned as they are for words with few neighbors in the adult mental lexicon. However, as children learn more and more words, neighborhoods for some words are expected to become denser. Therefore, orthographic representations of these words with a high neighborhood density are predicted to become more narrowly tuned as children grow. Castle et al. (1999) failed to find evidence for their prediction in their first study because Grade 6 students, as well as Grades 2 and 4 students, still showed form facilitation from nonword primes. However, in a subsequent longitudinal study, Castle et al. (2007) found that children who showed facilitation from orthographic neighbor (word) primes when they were in Grade 3 no longer showed significant form facilitation when they took part in the same experiment after 2 years[6].

A similar argument was put forward based on interaction-based models. Andrews and her colleagues (Andrews & Hersch; Andrews & Lo, 2012) found that college students with better spelling knowledge showed stronger inhibition from orthographic neighbor primes than other college students with poorer spelling knowledge. They then claimed that form inhibition was discovered only in good spellers because only when orthographic representations are fully specified can the masked word prime be successfully identified within a short prime duration and simultaneously activate its neighbors to compete for selection. Even though this argument is less direct than the lexical tuning hypothesis about the developmental nature of lexical representations, it agrees with the idea that the patterns of form priming are modulated by the quality of orthographic representations.

2.1.2. Form Priming Experiments in the L2

The previous section summarized the findings of research on L1 word recognition, and this section will review what similar and different findings have been obtained in L2 word identification research. Given the purpose of the study that was presented in Section 1.2, similarities between L1 and L2 lexical processing will support the hypothesis that both L1 and

---

[6] Baxter, Droop, Van Den Hurk, Bekkering, Dijkstra, and Léoné (2021) report similar results obtained in their vocabulary training study with early L2 learners. In this study, participants learned L2 words (e.g., *beak* and *tire*) better when the target L2 words were presented with orthographically (e.g., *beak, bead, beam*) or semantically similar words (e.g., *tire, brake, gear*) during the training phase compared to when the target L2 words were presented with orthographically (e.g., *beak, lawn, poison*) or semantically dissimilar words (e.g., *tire, leaf, plumber*). Baxter and her colleagues claimed that when similar words were presented, participants might have paid greater attention to the lexical dimension in which the similarity occurs (i.e., orthography or semantics), such that they could build more precise lexical representations. This argument, which was made in L2 research, is consistent in a broader sense with the lexical tuning hypothesis suggested in L1 research.

L2 operate based on the same mechanism, and differences will motivate the research questions of the present study.

Table 3 illustrates the results of previous form priming experiments that have been conducted using L2 words. For easier comparison, the results of some L1 studies are also summarized in Table 3. In these L1 studies, the same stimuli that were used for L2 participants were presented to L1 participants.

Several L2 studies show patterns of priming that were observed in L1 speakers. Bijeljac-Babic, Biardeau, and Grainger (1997) report that proficient French-English bilinguals showed form inhibition from L2 word primes, although beginning bilinguals showed a null effect. Elgort and her colleagues explored how the patterns of form priming changed if pseudowords were trained under intentional (Elgort, 2011) and incidental learning conditions (Elgort & Warren, 2014). Elgort (2011) found that, in an unmasked priming experiment, trained pseudoword primes produced a null priming effect whereas untrained nonword primes produced facilitation. Similarly, Elgort and Warren (2014) found that incidentally learned pseudoword primes yielded a null priming effect unlike untrained nonword primes that produced facilitation. In short, these L2 studies showed either significant inhibition from the L2 word prime (Bijeljac-Babic et al., 1997) or the PLE, i.e., different patterns of form priming that are produced by trained pseudoword primes (a null priming effect) and untrained nonword primes (facilitation) (Elgort, 2011; Elgort & Warren, 2014).

However, subsequent studies did not show similar results. Qiao and Forster (2017) observed facilitation from the word prime in Chinese L2 learners of English. After training pseudowords to the same participants, these researchers could not detect the PLE, unlike in an L1 word training study that used the same materials (Qiao & Forster, 2013) or in other L2 word training studies (Elgort, 2011; Elgort & Warren, 2014). Specifically speaking, in the Qiao and Forster study (2017), trained L2 pseudoword form primes still produced facilitation just like nonword form primes. In addition, in subsequent studies (Jiang, 2021; Nakayama & Lupker, 2018) it was found that L2 word form primes produced facilitation even in proficient L2 speakers and with L2 word stimuli that were supposed to be familiar to the participants.

At the same time, facilitation from similar-looking L2 word primes has also been discovered in another body of research that investigated morphological processing during recognition of complex words (Diependaele, Duñabeitia, Morris, & Keuleers, 2011; Heyer & Clashen, 2015; Li, Jiang, & Gor, 2017; Li & Taft, 2020; Li, Taft, & Xu, 2017). These studies all included (a) an orthographic control condition in which the related prime had neither any morphological nor semantic relationship with the target (e.g., *freeze-free*) in addition to one or two critical condition(s): (b) the orthographically and (seemingly) morphologically related, but semantically opaque (e.g., *corner-corn*) condition and/(c) the orthographically, morphologically and semantically all related (e.g., *viewer-view*) condition. In these studies, L1 speakers' lexical decisions were usually facilitated after encountering related primes under the critical condition(s) but not under the control condition. These findings indicate that in native visual word recognition of complex words, morphology as well as semantics play an important role, and automatic grammatical processing (morphological decomposition) and facilitative semantic priming takes place. Likewise, L2 speakers' lexical decisions were usually facilitated by the related prime under the critical condition(s), which appears to indicate that they can also decompose morphologically complex words. However, L2 speakers also showed facilitation even under the control condition, posing a challenge to the conclusion that morphological priming in the L2 is purely morphological in nature (e.g., Heyer & Clashen, 2015). This is because the facilitation found under the three conditions with L2 speakers in the morphological processing studies (Diependaele et al., 2011; Heyer & Clashen, 2015; Li et al., 2017; Li & Taft, 2020; Li et al., 2017) could all be due to facilitative priming from form overlap. This possibility is supported by De Moor and Brysbaert (2000) who showed that form priming is not just limited to prime-target pairs with the same length (e.g., Dutch words *buik-BUIL*) but extends to pairs with different lengths (e.g., *laars-AARS*).

To sum up, although indicators for lexical competition in the L2 (form inhibition or a null form priming effect) have sometimes been detected, evidence of the absence of, or a very weak lexical competition (form facilitation) has been more frequently observed. Moreover, although one study found a significant inhibitory neighbor priming effect in the L2 (Bijeljac-Babic et al., 1997), this study recruited same-script (Dutch-English) bilinguals. Similarly, in two studies that reported a null orthographic priming effect for newly trained L2 pseudowords (Elgort, 2011; Elgort & Warren, 2014), participants' L1 was not controlled, so some participants' L1 (e.g., French or German) used the same script as the L2 English. In contrast, when participants' L1 used a Chinese or Japanese script rather than the Roman alphabet, form facilitation has always been observed. Therefore, Nakayama and Lupker (2018) suspected that there might be no lexical competition in different-script bilinguals' visual L2 word recognition.

Table 3

The Results of Unmasked Orthographic Form Priming Experiments in the L2 and the L1 (which Used the Same Materials as Those in the L2 Experiments)

L1/L2	Study	Prime lexicality	Prime duration	Prime length	Target lexicality	Target duration	Target N density	RT: Related	RT: Unrelated	Unrelated - related	Special manipulations	Language
L2	Elgort (2011)	Word	522	7~9	Word	522	Not reported	753	739	-14	Known words	English
L2	Elgort (2011)	Nonword	522	7~9	Word	522	Not reported	678	739	61*		English
L2	Elgort (2011)	Word	522	7~9	Word	522	Not reported	760	780	20	Newly trained words	English
L2	Elgort (2011)	Nonword	522	7~9	Word	522	Not reported	705	780	75*		English
L2	Nakayama & Lupker (2018) Experiment 6	Word	175	4~5	Word	Until being responded	5.5	743	735	-8	No stats.	English
L2	Nakayama & Lupker (2018) Experiment 6	Nonword	175	4~5	Word	Until being responded	5.5	710	741	31		English
L2	Nakayama & Lupker (2018) Experiment 6	Word	175	4~5	Nonword	Until being responded	5.1	819	824	5		English
L2	Nakayama & Lupker (2018) Experiment 6	Nonword	175	4~5	Nonword	Until being responded	5.1	800	816	16		English
L1	Park (2021) Experiment 2	Word (Low F)	Until being responded	6~10	Word	Until being responded	1.7	630	633	3	Mid-term priming method.	English
L1	Park (2021) Experiment 2	Word (High F)	Until being responded	6~10	Word	Until being responded	1.7	631	621	-10		English
L1	Park (2021) Experiment 2	Word	Until being responded	6~10	Nonword	Until being responded	1	666	664	-2		English
L2	Park (2021) Experiment 2	Word (Low F)	Until being responded	6~10	Word	Until being responded	1.7	783	833	50*		English
L2	Park (2021) Experiment 2	Word (High F)	Until being responded	6~10	Word	Until being responded	1.7	812	851	40*		English
L2	Park (2021) Experiment 2	Nonword	Until being responded	6~10	Nonword	Until being responded	1	846	840	-6		English

F: Frequency, No stats.: Reporting main effects only without statistical analysis for the form priming effect under each condition, *: $p < .05$.

Table 4

The Results of Unmasked Orthographic Form Priming Experiments in the L2 and the L1 (which Used the Same Materials as Those in the L2 Experiments)

L1/L2	Study	Prime lexicality	Prime duration	Prime length	Target lexicality	Target duration	Target N density	RT: Related	RT: Unrelated	Unrelated - related	Special manipulations	Language
L1→L2	Bijeljac-Babic et al. (1997) Experiment 1	Word (L1)	57	4	Word (L2)	Until being responded	Not reported	732	704	-28	A significant main effect (inhibition)	English
L2	Bijeljac-Babic et al. (1997) Experiment 1	Word (L2)	57	4	Word (L2)	Until being responded	Not reported	688	634	-54		English
L1	Bijeljac-Babic et al. (1997) Experiment 2	Word (L1)	57	4	Word (L1)	Until being responded	Not reported	740	690	-50	Monolingual participants, no stats.	French
L2→L1	Bijeljac-Babic et al. (1997) Experiment 2	Word (L2)	57	4	Word (L1)	Until being responded	Not reported	734	730	-4		French
L1	Bijeljac-Babic et al. (1997) Experiment 2	Word (L1)	57	4	Word (L1)	Until being responded	Not reported	785	734	-51	Low-prof. bilingual participants, no stats.	French
L2→L1	Bijeljac-Babic et al. (1997) Experiment 2	Word (L2)	57	4	Word (L1)	Until being responded	Not reported	769	752	-17	Low-prof. bilingual participants	French
L1	Bijeljac-Babic et al. (1997) Experiment 2	Word (L1)	57	4	Word (L1)	Until being responded	Not reported	757	729	-28	high-prof. bilingual participants, no stats.	French
L2→L1	Bijeljac-Babic et al. (1997) Experiment 2	Word (L2)	57	4	Word (L1)	Until being responded	Not reported	792	749	-43*	High-prof. bilingual participants	French
L1	Qiao & Forster (2012) Experiment 1	Word (Trained)	50	6~8	Word	500	1.1	544	567	23	Newly trained word primes	English
L1	Qiao & Forster (2012) Experiment 1	Nonword	50	6~8	Word	500	1.1	537	566	29*	Untrained nonword primes	English
L1	Qiao & Forster (2012) Experiment 2	Word	50	6~8	Word	500	1.1	642	664	22*	Newly trained word targets (Session 1)	English
L1	Qiao & Forster (2012) Experiment 2	Word	50	6~8	Word	500	1.1	610	645	35*	Newly trained word targets (Session 2)	English
L1	Qiao & Forster (2012) Experiment 2	Word	50	6~8	Word	500	1.1	580	620	40*	Newly trained word targets (Session 3)	English
L1	Qiao & Forster (2012) Experiment 2	Word	50	6~8	Word	500	1.1	587	598	11	Known word targets (Session 1)	English
L1	Qiao & Forster (2012) Experiment 2	Word	50	6~8	Word	500	1.1	571	570	-1	Known word targets (Session 2)	English
L1	Qiao & Forster (2012) Experiment 2	Word	50	6~8	Word	500	1.1	559	558	-1	Known word targets (Session 3)	English
L1	Qiao & Forster (2012) Experiment 2	Word	50	6~8	Word	500	1.1	612	618	6	Newly trained word primes	English
L1	Qiao & Forster (2012) Experiment 2	Nonword	50	6~8	Word	500	1.1	508	536	28*	Untrained nonword primes	English
L2	Elgort & Warren (2014)	Word	56	5~6	Word	500	1.5	907	910	3	Primes were newly trained words.	English
L2	Elgort & Warren (2014)	Nonword	56	5~6	Word	500	1.5	887	910	23*		English
L2	Qiao & Forster (2017)	Word	50	6~8	Word	500	Not reported	804	841	37*	Newly trained word targets (Session 1)	English
L2	Qiao & Forster (2017)	Word	50	6~8	Word	500	Not reported	778	823	45*	Newly trained word targets (Session 2)	English
L2	Qiao & Forster (2017)	Word	50	6~8	Word	500	Not reported	732	790	58*	Newly trained word targets (Session 3)	English
L2	Qiao & Forster (2017)	Word	50	6~8	Word	500	Not reported	765	826	61*	Known word targets (Session 1)	English
L2	Qiao & Forster (2017)	Word	50	6~8	Word	500	Not reported	822	879	57*	Known word targets (Session 2)	English
L2	Qiao & Forster (2017)	Word	50	6~8	Word	500	Not reported	835	879	44*	Known word targets (Session 3)	English
L2	Qiao & Forster (2017)	Word	50	6~8	Word	500	Not reported	809	862	53*	Newly trained word primes	English
L2	Qiao & Forster (2017)	Nonword	50	6~8	Word	500	Not reported	842	875	33*	Untrained nonword primes	English

L1/L2	Study	Prime lexicality	Prime duration	Prime length	Target lexicality	Target duration	Target N density	RT: Related	RT: Unrelated	Unrelated - related	Special manipulations	Language
L1	Nakayama & Lupker (2018) Experiment 2	Word	67	4~5	Word	Until being responded	5.5	582	560	-22*		English
L1	Nakayama & Lupker (2018) Experiment 2	Nonword	67	4~5	Word	Until being responded	5.5	568	558	-10		English
L1	Nakayama & Lupker (2018) Experiment 2	Word	67	4~5	Nonword	Until being responded	5.1	664	630	-34*		English
L1	Nakayama & Lupker (2018) Experiment 2	Nonword	67	4~5	Nonword	Until being responded	5.1	627	640	13		English
L2	Nakayama & Lupker (2018) Experiment 1	Word	67	4~5	Word	Until being responded	5.5	704	725	21	No stats.	English
L2	Nakayama & Lupker (2018) Experiment 1	Nonword	67	4~5	Word	Until being responded	5.5	704	729	25		English
L2	Nakayama & Lupker (2018) Experiment 1	Word	67	4~5	Nonword	Until being responded	5.1	793	800	7		English
L2	Nakayama & Lupker (2018) Experiment 1	Nonword	67	4~5	Nonword	Until being responded	5.1	781	803	22		English
L2	Nakayama & Lupker (2018) Experiment 4	Word	67	4~5	Word	Until being responded	8.7	665	692	27*		English
L2	Nakayama & Lupker (2018) Experiment 5	Word	67	4~5	Word	Until being responded	5.5	644	670	26	No stats.	English
L2	Nakayama & Lupker (2018) Experiment 5	Nonword	67	4~5	Word	Until being responded	5.5	655	667	12		English
L2	Nakayama & Lupker (2018) Experiment 5	Word	67	4~5	Nonword	Until being responded	5.1	695	718	23		English
L2	Nakayama & Lupker (2018) Experiment 5	Nonword	67	4~5	Nonword	Until being responded	5.1	688	717	29		English
L2	Jiang (2021)	Word	50	5~7	Word	Until being responded	2.3	549	560	11		English
L2	Jiang (2021)	Word	50	5~7	Word	Until being responded	2.3	821	884	64*		English
L1	Park (2021) Experiment 1	Word (Low F)	67	4~5	Word	Until being responded	9.2	699	663	-36*		English
L1	Park (2021) Experiment 1	Word (High F)	67	4~5	Word	Until being responded	9.6	689	651	-38*		English
L1	Park (2021) Experiment 1	Word (Low F)	67	4~5	Nonword	Until being responded	6.6	754	753	-1		English
L1	Park (2021) Experiment 1	Word (High F)	67	4~5	Nonword	Until being responded	6.3	732	747	15		English
L2	Park (2021) Experiment 1	Word (Low F)	67	4~5	Word	Until being responded	9.2	852	856	4		English
L2	Park (2021) Experiment 1	Word (High F)	67	4~5	Word	Until being responded	9.6	837	831	-6		English
L2	Park (2021) Experiment 1	Word (Low F)	67	4~5	Nonword	Until being responded	6.6	924	928	4		English
L2	Park (2021) Experiment 1	Word (High F)	67	4~5	Nonword	Until being responded	6.3	913	933	20*		English

F: Frequency, No stats.: Reporting main effects only without statistical analysis for the form priming effect under each condition, *: $p < .05$.

2.1.3. The Reason for Form Facilitation from the Word Prime in the L2

Given the reports indicating weak form inhibition from the word prime in the L2, Qiao and Forster (2017) claimed that this is because L2 words are not stored in lexical memory where L1 words are stored. This memory-based account is in the same line with the hypothesis that L1 words are stored in episodic memory at an initial stage of vocabulary acquisition, and there is no lexical competition between a newly learned word and its orthographic neighbor while these learned words are represented in episodic memory (Gaskell & Dumay, 2003; Lindsay & Gaskell, 2010; McClelland, McNaughton, & O'Reilly, 1995; Tamminen & Gaskell, 2013). Qiao and Forster (2017), however, argued that unlike acquired L1 words that eventually transfer to lexical memory after consolidation, L2 words are represented in episodic memory for a more extended period. They then assumed that neither lexical tuning nor lexical competition takes place in episodic memory.

Other researchers have suggested that weak inhibition for L2 word prime-target pairs is because even though L2 words are stored in lexical memory, their representations are not robust enough to produce form inhibition (Gor & Cook, 2020; Gor, forthcoming; Gor et al., 2021; Jiang, 2021). This linguistic representation-based account was examined in Park's study (2021).

Park (2021) first used the medium-term priming method with long word stimuli (6~10 letters) where the prime was consciously identifiable, but the relationship between the prime and the target was not obvious. In this experiment, Korean late L2 learners of English showed significant form facilitation whereas L1 speakers showed a null effect. This finding appeared to support the idea that the L2 lexicon behaves differently from the L1 lexicon. However, when unknown and unfamiliar L2 stimuli (that were identified by a post-experiment survey) were excluded from analysis, high-frequency primes did not lead to faster responses to the target, just like L1 word primes, although low-frequency primes still elicited significantly faster lexical decisions. These findings were consistent with those observed in phonological priming auditory word identification experiments that showed the effects of familiarity with L2 words and of prime frequency on lexical competition in the L2 (Cook & Gor, 2015; Gor & Cook, 2020). Furthermore, when the effect of participants' spelling knowledge – which was argued to be an indicator for the quality of orthographic representations (Andrews & Hersch, 2010, Andrews & Lo, 2012; Andrews, Veldre, & Clarke, 2020) – on response latencies to well-known L2 words was examined, it was found that, whereas high-frequency primes produced non-significant inhibition regardless of participants' spelling test scores, the patterns of form priming from low-frequency primes were modulated by participants' spelling test scores. Specifically speaking, while poor L2 spellers showed form facilitation from the low-frequency primes, good L2 spellers showed a null effect.

Park (2021) also used the masked priming paradigm to explore the nature of masked orthographic priming for L2 words. As shown in Tables 3 and 4 above, the previous studies had suggested that when the prime is masked, it may be harder to observe a nativelike pattern of form priming in the L2 relative to when the prime is unmasked. The results of reaction time analysis in Park's priming experiment (2021) showed that the word form prime inhibited L1 speakers' lexical decisions on the target, whereas it produced a null priming effect in L2 speakers. Since Nakayama and Lupker (2018) found only form facilitation in the L2 in a series of form priming experiments where proficient Japanese L2 learners of English were recruited who even showed significant facilitative L2-L1 translation priming[7], the null effect observed in Park's study (2021) was interpreted to be in conflict with Nakayama and Lupker's (2018) assumption that there is no lexical competition in different-script bilinguals. However, unlike in the medium-term priming experiment, even after excluding unknown and unfamiliar items, L2 speakers still did not show a nativelike pattern of form priming (i.e., significant inhibition). In addition, contrary to expectations, L2 speakers' spelling knowledge did not modulate the pattern of form priming for well-known word stimuli in response latency analysis. Instead, when controlling for the effect of spelling knowledge, L2 speakers' proficiency significantly modulated the strength of form priming. In other words, whereas lower-proficiency L2 speakers showed form facilitation, high-proficiency L2 speakers did not. Considering the above-mentioned roles of neighborhood and vocabulary size in orthographic neighbor priming in L1 word recognition, the significant role of proficiency in determining the strength of orthographic priming effects is understandable if it is assumed that participants' proficiency might indicate L2 participants' neighborhood or vocabulary size. To be more specific, unlike in Park's medium-term priming experiment (2021) in which stimuli had few neighbors, stimuli had many neighbors in the masked priming experiment. This suggests that whereas the impact of orthographic precision on form priming was more easily observable when the neighborhood size of stimuli was controlled, it might have been less prominent when stimuli had many neighbors since the neighborhood size also influences orthographic priming. Therefore, although the effect of spelling knowledge on form facilitation was non-significant, the finding that proficiency, which might be related to L2

[7] The L2-L1 translation priming effect found in the L2 participants in the Nakayama and Lupker (2018) study suggests that those participants were proficient enough to be able to activate semantic information of the masked L2 word prime, which has been known to require high proficiency (Nakayama, Ida, & Lupker, 2016).

participants' vocabulary size, influences orthographic priming still signifies the developmental nature of L2 orthographic representations.

The above-mentioned discoveries from the two experiments in Park's study (2021) support the linguistic representation-based account which posits that weak form inhibition in the L2 is due to less-specified or weak L2 lexical representations (Gor, forthcoming; Gor et al., forthcoming, Jiang, 2021). However, one may wonder how L2 orthographic representations, which would not be influenced by the obstacles for L2 phonological encoding (see Section 1.1.2) could also be fuzzy. Several plausible explanations can be drawn from previous research.

The first explanation can be found from Perfetti's lexical quality hypothesis. Perfetti (1992, 2007) proposed that with experience of reading, L1 readers implicitly learn many grapheme-phoneme conversion rules. In other words, while encountering specific letter strings numerous times, L1 readers get used to phonologically decoding them, and this redundant knowledge of sub-lexical components contributes to the establishment and refinement of orthographic representations (see also Welcome & Trammel, 2017). However, late different-script bilinguals' exposure to Roman alphabets is far more limited than that of adult L1 speakers. Therefore, low levels of familiarity with Roman alphabets could be a reason for late L2 speakers' inefficient orthographic encoding. As noted earlier, when L1 and L2 speakers were trained with the same pseudowords under identical environments, only L1 speakers showed a null effect in a priming experiment with these pseudoword primes, whereas L2 speakers showed form facilitation (Qiao & Forster, 2013, 2017). This may indicate L2 speakers' reduced ability for orthographic encoding. Interestingly, L2 vocabulary learning studies that showed reduced facilitation from pseudoword orthographic neighbor primes in an LDT after training (Elgort, 2011; Elgort & Warren, 2014) included early L2 learner participants. In addition, Elgort and

Warren (2014) discovered that the strength of form facilitation from newly learned pseudoword primes was negatively moderated by participants' AOAs. These findings appear to suggest that if L2 learners are exposed to L2 script from their early childhood and get used to it, they can also efficiently encode the orthographic forms of novel words like L1 speakers.

The way that late L2 speakers learn vocabulary may also influence the establishment of weak orthographic representations. L1 speakers usually establish phonological representations first. For that reason, when L1 speakers learn the spelling of L1 words, these words are often already familiar to them in terms of the pronunciation and meaning. Therefore, L1 speakers would only need to remember the orthographic form of the word, which may be easily acquired based on their already existing phoneme-grapheme conversion rules. However, when L2 speakers learn vocabulary, they must usually memorize pronunciation, spelling and meaning at the same time (Bordag et at., 2021; Gor et al., 2021; Jiang, 2000; Jiang & Pae, 2021). Therefore, the orthographic encoding of L2 words would be much more effortful than that for L1 words, and the quality of L2 orthographic representations would also be poor at an initial stage of lexical development.

The memory-based account (Qiao & Forster, 2017) is not supported by the findings from Park's study (2021). According to this account, there should be form facilitation for L2 word prime-target pairs because episodic memory is assumed to be structured somewhat differently from lexical memory such that there are no competitive processes among episodic memory representations, at least if the prime is invisible. However, L2 participants in Park's study (2021) showed a null priming effect in the masked priming experiment. This finding indicates that although it was weaker compared to that observed in L1 speakers, a certain degree of lexical competition between L2 words might have occurred. Thus, this finding does not support the assumption that there is no lexical competition between episodic representations.

Another challenge for the memory-based account is the fact that form inhibition from masked L2 word primes was sometimes observed. Bijeljac-Barbic and her colleagues (Bijeljac - Barbic et al., 1997) found that their L2 participants showed inhibition for the L2 word target when it was primed by its L2 orthographic neighbor, so explaining this finding with the memory-based account is problematic. This problem may be avoided if it is assumed that L2 words can also transfer from episodic memory to lexical memory just like L1 words if the L2 lexical representations are consolidated through repeated exposure. (Nakayama, Ida, & Lupker, 2016). If so, the reason for the form inhibition observed in the previous study (Bijeljac-Barbic et al., 1997) could be considered due to the high L2 proficiency of the participants in that study. However, Nakayama and Lupker (2018) reported that for L2 participants who showed a facilitative cross-language (L2→L1), the masked translation priming effect still showed form facilitation in the L2. The episodic L2 hypothesis (Jiang & Forster, 2001; Qiao & Forster, 2012) supposes that the reason why an L2-L1 translation priming effect is not frequently observed is that L2 words are stored in episodic memory contrary to L1 words that are represented in lexical memory, and that priming from episodic memory to lexical memory is restricted if the prime is accessed subconsciously. Thus, according to this hypothesis, the L2-L1 translation priming detected in the Nakayama and Lupker study (2018) should indicate that the participants' proficiency level was high enough to store at least some L2 words in lexical memory (see Nakayama et al., 2016). Nevertheless, these participants did not show form inhibition from the same orthographic L2 word primes which produced facilitation when they had been used as translation primes. Therefore, in order for the memory-based account to be supported, an explanation is necessary for how the L2 words that failed to yield lexical competition due to their representations in episodic memory could have produced the L2-L1 translation priming effect shown in Nakayama and Lupker's study (2018).

To summarize, many form priming studies have shown robust form facilitation from the L2 word prime, particularly in late different-script L2 speakers, and this pattern of form priming contrasts with the patterns that have been observed in L1 speakers (a null form priming effect or inhibition, see Table 1). Thus, to investigate the reason for this phenomenon, Park (2021) conducted a study and obtained several findings that suggest that form facilitation in L2 may occur because of a poor quality of L2 orthographic representations and a small neighborhood size of the prime. However, making strong claims for the FLR hypothesis may still be limited for several reasons. First, unlike in the unmasked priming experiment, both low- and high-frequency primes produced a null effect in response latency analysis of the masked priming LDT, even after excluding unfamiliar stimuli from the analysis. For that reason, one may want to see further evidence that more clearly demonstrates the developmental nature of L2 lexical representations (such as significant form inhibition in high-proficiency L2 speakers that contrasts with significant facilitation in low-proficiency L2 speakers). Second, although Park recruited only top scorers on TOEIC (Test of English for International Communication)[8], these participants still did not show a native-like pattern of priming (inhibition), even when the most frequent English words were used as the prime. Therefore, evidence has not yet been found for the assumption that form inhibition would eventually be observed (even when the prime is masked) if L2 orthographic representations become fully specified, even in different-script L2 speakers.

[8] TOEIC is a paper-and-pencil, multiple-choice standardized English proficiency test that measures non-native speakers' reading and listening comprehension skills. This test is widely used, particularly in East Asia.

Therefore, in order to overcome these limitations in Park's study (2021), further research is required.

2.2. Nonnative Phonological Categories and Visual Word Processing

2.2.1. The Impact of Fuzzy L2 Word Form Representations on Visual Lexical Access

As noted in Section 1.1.2, acquisition of difficult L2 phonological contrasts that are not discriminated in the L1 is known to be very difficult. The L2 speaker's insensitivity to nonnative contrasts found in the reviewed previous studies (Broersma, 2012; Pallier et al., 2001; Weber & Cutler, 2004) might occur because phonological representations of words with confusing contrasts (e.g., *flash-flesh*) are not separate in the L2 speaker's mental lexicon. Put differently, based on the hypothesis that mental representations of phonetic segments play a crucial role in accurate sound perception (Flege, 2003; Meador et al., 2003; Thorn & Gathercole, 2001), it can be reasoned that L2 speakers' limited ability to differentiate difficult phonological contrasts is not only the reason for less-precise L2 phonological representations but also the outcome of these coarsely-specified representations. However, since the previous studies were conducted using auditory stimuli (Broersma, 2012; Pallier et al., 2001; Weber & Cutler, 2004), they were limited in convincingly claiming that L2 speakers' non-nativelike processing is due to representations as well as misperception. Thus, to avoid confusion between representation and perception, Ota, Hartsuiker and Haywood (2009) investigated the impact of less-precise L2 word form representations on L2 processing using a visual task that did not involve auditory perception.

The task that Ota et al. (2009) used was a semantic-relatedness judgment task, in which participants had to judge whether two visually presented words (e.g., *KEY-LOCK*) were semantically related. Since the /l/-/r/ contrast does not exist as a phonological contrast in Japanese, the researchers predicted that Japanese L2 learners of English would confuse *LOCK* with *ROCK*, such that they would often respond that *KEY-ROCK* were related. The results revealed that unlike L1 speakers, Japanese L2 participants made more errors on, and responded slower to, word pairs in which one member of the semantically related pairs was replaced by its near-homophone (e.g., *KEY-ROCK* which was altered from *KEY-LOCK*). On the other hand, Arabic L2 participants whose native language has separate L1 phonemes related to /l/-/r/ but not for /b/-/p/, did not show confusion on these critical pairs. Instead, they were confused when encountering pairs where a word with /b/ was replaced by its minimal pair with /p/ (e.g., *SAND-PEACH*, *PEACH* is a near-homophone of *BEACH*). In contrast, Japanese L2 participants whose L1 discriminates relevant sounds related to the /b/-/p/ contrast did not show inhibition when responding to these latter pairs. The findings of Ota and his colleagues were then replicated in their subsequent study (Ota, Hartsuiker, & Haywood, 2010). In this second study, the researchers found that interference from near-homophones is not just limited to words containing the /l/-/r/ contrast but also occurs for words with other nonnative contrasts (/æ/–/ʌ/ and /b/-/v/) during Japanese-English bilinguals' visual word processing.

The findings from Ota's studies (Ota et al., 2009, 2010) can be interpreted in two different ways. First, the form confusion that was observed for *rock* and *lock* could be considered to have occurred at the level of phonology. This explanation is in line with the interpretation of similar findings observed in L1 research. For instance, Van Orden (1987) reported that that L1 speakers found it difficult to decide that *rows* (a homophone with *rose*) is not a kind of flower in a semantic categorization task. Van Orden (1987) claimed that the finding of his study indicates automatic access to the phonological information of written words. Following this logic, Ota and his colleagues (Ota et al., 2009, 2010) thought that the confusion of *lock* with *rock* occurred due to automatic activation of the phonology of the two words that were not easily discriminated.

Another explanation is possible when considering the difficulty that L2 learners experience at the stage of orthographic encoding. Specifically, it can be assumed that nonnative phonological contrasts pose an extra challenge for orthographic encoding because the similar phonological forms of two words with a nonnative contrast interfere with accurate orthography-meaning mappings (Bordag et al., 2021; Gor et al., 2021; Gor, forthcomding) as L2 homophones (e.g., *tale-tail*) do. In this case, the fuzziness present in phonological representations prevents the development of precise orthographic representations such that orthographic representations of words with confusing contrasts are fuzzier than those without a difficult contrast. In other words, what produces form confusion for *lock* and *rock* is the fuzzy orthographic representations of these two words rather than the fuzzy phonological representations, even though the fundamental reason for unsuccessful orthographic encoding is the phonology.

No matter which explanation (i.e., phonology or orthography) captures the exact mechanism of the form confusion for *lock* and *rock*, Ota and his colleagues' discoveries (2009, 2010) clearly demonstrate that form confusion for two words with a nonnative contrast occurs even in a written task, and this demonstrates the fuzzy form representations of L2 words with a confusing contrast. At the same time, their discoveries lead to the hypothesis that the impact of difficult L2 phonological contrast (e.g., /l/-/r/) on visual word recognition would also be observable in a masked orthographic priming experiment using L2 words. The current study makes use of this possibility to explore how fuzzy L2 form representations influence the identification of written L2 words. The following section reviews what findings have been made in L1 research that used the same research method that the current study uses.

2.2.2. Masked Phonological Priming in the L1

The terminology "phonological priming" has been used a little differently in two different bodies of the literature. In spoken word identification research, the phonological prime refers to a stimulus whose several phonemes overlap with the target. The pattern of priming that has received more attention from researchers is inhibition, which is produced by the phonological prime sharing the initial three phonemes with the target (see Gor, 2018 for a review). On the other hand, in written word recognition research, the phonological prime refers to a homophone or pseudo-homophone of the target, and this prime is believed to produce facilitation if the effect of the lexical competition is controlled (see Rastle & Brysbaert, 2006 for a review). This section reviews previous studies on masked phonological priming used in written word identification research.

In one of the most frequently cited early masked homophone priming experiments, Humphreys, Evett and Taylor (1982) briefly presented participants with both the masked prime and masked target (25~50 ms) and let them type the target word. In this perceptual identification task, participants significantly more correctly identified the target word (e.g., *MADE*) if the prime was its homophone (e.g., *maid*) than when the prime was its grapheme control (e.g., *mark*). Facilitative phonological priming has also been observed in the word naming task in which participants are asked to read aloud the single target word presented on the screen. For instance, Bowers, Vigliocco and Haan (1998) showed that participants responded faster when the target (e.g., *SAIL*) followed its homophone prime (e.g., *sale*) compared to when it was preceded by an unrelated prime (e.g., *butt*). A comprehensive review of masked phonological experiments (Rastle & Brysbaert, 2006) reports robust facilitative priming both in the perceptual identification task and the naming task. Strong phonological priming in the word naming task in particular has been thought to occur because the phonological information of the prime is more strongly activated when the task requires a spoken response (e.g., Kinoshita & Norris, 2012b; Rastle & Brysbaert, 2006).

As shown above, although various tasks such as perceptual identification or word naming have been used to investigate masked phonological priming effects, some researchers claim that the LDT is the best task for examining the role of phonology during lexical access because other tasks involve other cognitive processes such as offline guessing in perceptual identification or preparation of a speech response in word naming (Andrews, 1997; Davis, Castles, & Iakovidis, 1998; Kinoshita & Norris, 2012a; Rastle & Brysbaert, 2006).

Table 5 presents phonological priming effects observed in previous studies that used the LDT. The results of a few word naming experiments are also included for comparison. Grainger and Ferrand (1994) reported that facilitative masked homophone priming (e.g., *real-REEL*) was observed in their study. However, subsequent studies (Bowers et al., 1998; Shen & Forster, 1999[9]) report that, unlike in word naming, it is not that easy to observe facilitation from homophone primes in LDTs. This is because although the pre-activation of sub-lexical components (i.e., phonemes) of the target through the masked homophone prime may facilitate access to the target, but if the prime activates its orthographic neighbors or phonological competitors, it can also produce form inhibition.

Therefore, in order to explore the phonological priming effect under a condition where lexical competition is reduced, pseudo-homophones (nonwords) have been used as the prime

[9] Shen and Forster (1999) claimed that the null priming from homophones in their study was because native Chinese speakers whose L1 use a logographic script do not depend on the pronunciation of words for lexical access.

more frequently than homophones because it is known that nonword primes are likely to trigger weaker lexical competition than word primes (see Section 2.1.1). When the priming effects from pseudo-homophone primes and orthographic primes were compared, Ferrand & Grainger (1992) found greater facilitation from pseudo-homophone primes over grapheme control primes in French. However, findings of subsequent studies suggest that phonological priming may be fragile (Bowers et al., 1998; Davis et al., 1998; Kinoshita & Norris, 2012a; Rastle & Brysbaert, 2006), probably because pseudoword primes can also trigger lexical competition by activating the neighborhood of the target.

To conclude, although previous research indicates that the phonological (homophone or pseudo-homophone) prime can facilitate target recognition through activation of phonological codes of the prime even when no sound is played or when articulation of the prime is not required, facilitative masked phonological priming effects have been difficult to detect at least in LDTs. This has been interpreted to be because a phonological boost is canceled out by lexical competition at the level of orthographic representations. However, as reviewed in Section 1.1.3 and Section 2.1, previous research has found that lexical competition is weak in the L2. In addition, if bilinguals' L1 and L2 do not use the same script, the influence of L1 orthographic neighbors on the L2 target would be minimized (Bijeljac-Babic et al., 1997; Van Heuven et al., 1998). This reasoning supports the hypothesis that examining the masked orthographic priming effect on L2 word targets with difficult L2 phonological contrasts (e.g., *read-LEAD*) will contribute important insights to the investigation of L2 word recognition, particularly if L2 speakers' first and second languages use different scripts.

Before concluding this section, it is worth noting that although the previous research reviewed in this section has been conducted based on the idea that the form facilitation for homophone pairs comes from the early activation of the phonology of the target by the prime, this paper does not consider that it is the only reason for the expected form facilitation for *read-LEAD*. The next section describes all plausible reasons for this predicted pattern of priming.

2.2.3. Mechanisms of Form Facilitation for Prime-Target Pairs with Nonnative Contrasts

When expecting form facilitation for minimal pairs with non-native phonological contrasts, three possible mechanisms of this phenomenon are conceivable.

First, in consideration of the absence of a non-native phoneme (e.g., /l/ or /ɹ/) in Korean native speakers' L1 phoneme inventory, it is possible to imagine a situation where L2 speakers find it difficult to map the correct grapheme onto a non-native phoneme. Some researchers (e.g., Coltheart, 2006) claim that in order to represent lexical information in the mental lexicon, people depend more on the pronunciation than the visual form of words because their visual memory has a much lower capacity than their phonological memory. Therefore, if L2 learners are not sure about the exact pronunciation of a certain L2 phoneme, these learners could experience difficulty when trying to map a certain L2 grapheme onto this unfamiliar L2 phoneme. For example, Wang and Geva (2003) report that when Chinese ESL children who began to learn how to write at school were asked to write the spelling of English words including the letter string *th* (*teeth* and *thick*), many of them could not provide any letter for the phoneme /θ/, which does not exist in their L1. This finding illustrates the challenge for orthographic encoding of non-native phonemes. At the same time, this suggests that Korean L2 learners of English could have the same problems for orthographic encoding of non-native phonemes such as /l/ or /ɹ/ since they are not sure about which sound should be mapped onto the letter *l* or *r*. If this is the case, form facilitation for minimal pairs with confusing phonological contrasts in a priming experiment will originate from the level of fuzzy orthographic representations.

The second plausible mechanism is one in which letters for a nonnative phonological contrast are mapped onto the same L1 phoneme such that the prime and the target are confused. This mechanism is a little different from the first one in which the letters *l* or *r* find no sound on which they could be mapped. In the second scenario, letters are mapped onto a certain sound, but the same sound (or non-discriminable two sounds) for two letters is the source of the problem. Wang and Geva (2003) show this in a spelling task in which Chinese ESL learners often replaced the letter *sh* with *s*, and *th* with *s* or *z*. These researchers proposed that this phenomenon emerged because the participants adopted any L1 phoneme for non-native L2 phonemes when they represented L2-word form information in their memory. This L1 transfer may also occur during Korean L2 learners' vocabulary acquisition, such that both of the phonemes /l/ and /ɹ/ could be mapped onto the Korean phoneme "ㄹ". (1) This would result in form confusion between minimal pairs with a difficult phonological contrast. If so, form facilitation could be due to the wrong activation of the target (e.g., *lead*) when encountering the prime (e.g., *read*). Otherwise, (2) the single sound (or non-discriminable two sounds) for two letters could cause unsuccessful orthographic encoding such that the orthographic representations of words with nonnative contrasts (that are fuzzier that those without a difficult contrast) could produce stronger orthographic form facilitation. In either case, the origin of the facilitation is the orthography.

The third possible explanation can be provided based on the interpretation of facilitation for homophonous word pairs (e.g., *real-REEL* in Grainger & Ferrand, 1994) or pseudo-homophone prime-word target pairs (e.g., *pharm-FARM* in Rastle & Brysbaert, 2006) that has been observed in L1 word recognition research (see Section 2.2.2). In this scenario, the phonological code of the prime is activated when it is encountered, but this code still facilitates target identification because the pronunciation of the two words that are represented in the mental lexicon are homophonous. Put differently, facilitation could occur due to the early activation of the phonemes of the target by the prime.

Although it is difficult to identify the exact mechanism of form facilitation for *read-LEAD*, with the plausibility of the third mechanism in mind, it would be meaningful to look at what findings have been made when cross-language homophonous prime-target pairs are used. This is because cross-language phonological priming studies assume that form facilitation for cross-language homophone prime-target pairs is operated by the third mechanism described above. Therefore, before concluding this chapter, the present paper reviews this body of research.

Table 5

The Results of Masked Phonological Priming Experiments in the L1

Task	Study	Prime lexicality	Prime duration	Prime length	Target	Target duration	Target N density	RT: Related	RT: Unrelated	Unrelated - related	Special manipulations	Language
LDT	Ferrand & Grainger (1992) Experiment 1	Nonword (PSH)	64	4	Word (High F)	Until being responded	Not reported	582	605	23	No stats. A significant main effect of PSH primes over ORTH primes	French
LDT	Ferrand & Grainger (1992) Experiment 1	Nonword (ORTH)	64	4	Word (High F)	Until being responded	Not reported	601	605	4		French
LDT	Ferrand & Grainger (1992) Experiment 1	Nonword (PSH)	64	4	Word (Low F)	Until being responded	Not reported	607	644	37		French
LDT	Ferrand & Grainger (1992) Experiment 1	Nonword (ORTH)	64	4	Word (Low F)	Until being responded	Not reported	641	644	3		French
LDT	Ferrand & Grainger (1992) Experiment 1	Nonword (PSH)	64	4	Nonword	Until being responded	Not reported	710	697	-13		French
LDT	Ferrand & Grainger (1992) Experiment 1	Nonword (ORTH)	64	4	Nonword	Until being responded	Not reported	692	697	5		French
LDT	Ferrand & Grainger (1992) Experiment 2	Nonword (PSH)	64	4	Word (High F)	Until being responded	Not reported	587	603	16	With PSH nonword foils. No stats. Significant main effect of PSH primes over ORTH primes	French
LDT	Ferrand & Grainger (1992) Experiment 2	Nonword (ORTH)	64	4	Word (High F)	Until being responded	Not reported	608	603	-5		French
LDT	Ferrand & Grainger (1992) Experiment 2	Nonword (PSH)	64	4	Word (Low F)	Until being responded	Not reported	619	638	19		French
LDT	Ferrand & Grainger (1992) Experiment 2	Nonword (ORTH)	64	4	Word (Low F)	Until being responded	Not reported	640	638	-2		French
LDT	Ferrand & Grainger (1992) Experiment 2	Nonword (PSH)	64	4	Word (PSH)	Until being responded	Not reported	777	750	-27		French
LDT	Ferrand & Grainger (1992) Experiment 2	Nonword (ORTH)	64	4	Word (PSH)	Until being responded	Not reported	738	750	12		French
LDT	Ferrand & Grainger (1992) Experiment 2	Nonword (PSH)	64	4	Nonword	Until being responded	Not reported	701	704	3		French
LDT	Ferrand & Grainger (1992) Experiment 2	Nonword (ORTH)	64	4	Nonword	Until being responded	Not reported	680	704	24		French
LDT	Ferrand & Grainger (1992) Experiment 3	Nonword (PSH)	32	4	Word (High F)	Until being responded	Not reported	558	593	35	No stats. Significant main effect of PSH primes over control primes, but not over ORTH primes	French
LDT	Ferrand & Grainger (1992) Experiment 3	Nonword (ORTH)	32	4	Word (High F)	Until being responded	Not reported	563	593	30		French
LDT	Ferrand & Grainger (1992) Experiment 3	Nonword (PSH)	32	4	Word (Low F)	Until being responded	Not reported	599	620	21		French
LDT	Ferrand & Grainger (1992) Experiment 3	Nonword (ORTH)	32	4	Word (Low F)	Until being responded	Not reported	604	620	16		French
LDT	Ferrand & Grainger (1992) Experiment 3	Nonword (PSH)	32	4	Nonword	Until being responded	Not reported	656	633	-23*		French
LDT	Ferrand & Grainger (1992) Experiment 3	Nonword (ORTH)	32	4	Nonword	Until being responded	Not reported	638	633	-5		French

Task	Study	Prime lexicality	Prime duration	Prime length	Target	Target duration	Target N density	RT: Related	RT: Unrelated	Unrelated - related	Special manipulations	Language
LDT	Grainger & Ferrand (1994) Experiment 1	Word (High F, ORTH similar)	64	4~5	Word	Until being responded	Non reported	626	663	37*	• ORTH similar: The homophone prime that differed from the target by one letter, • ORTH dissimilar: The homophone prime that differed from the target by two letters	French
LDT	Grainger & Ferrand (1994) Experiment 1	Word (High F, ORTH dissimilar)	64	4~5	Word	Until being responded	Non reported	639	649	10		French
LDT	Grainger & Ferrand (1994) Experiment 1	Word (Low F, ORTH similar)	64	4~5	Word	Until being responded	Non reported	631	618	-13		French
LDT	Grainger & Ferrand (1994) Experiment 1	Word (Low F, ORTH dissimilar)	64	4~5	Word	Until being responded	Non reported	625	611	-14		French
LDT	Grainger & Ferrand (1994) Experiment 2	Word (Homophone)	64	4~5	Word	Until being responded	Non reported	624	651	27	Without PSH nonword foils. No Stats.	French
LDT	Grainger & Ferrand (1994) Experiment 2	Word (ORTH)	64	4~5	Word	Until being responded	Non reported	671	640	-31		French
LDT	Grainger & Ferrand (1994) Experiment 2	Word (Homophone)	64	4~5	Word	Until being responded	Non reported	746	718	-28	With PSH nonword foils. No stats.	French
LDT	Grainger & Ferrand (1994) Experiment 2	Word (ORTH)	64	4~5	Word	Until being responded	Non reported	748	717	-31		French
LDT	Grainger & Ferrand (1994) Experiment 3	Word (Homophone)	64	3~6	Word	Until being responded	Non reported	617	649	32		English
LDT	Grainger & Ferrand (1994) Experiment 3	Word (ORTH)	64	3~6	Word	Until being responded	Non reported	670	647	-23		English
Naming	Bowers et al. (1998) Experiment 5	Word (High F, Homophone)	50	3~6	Word (Low F)	500	Non reported	504	525	21	No stats. Significant main effect	English
Naming	Bowers et al. (1998) Experiment 5	Word (Low F, Homophone)	50	3~6	Word (High F)	500	Non reported	487	520	33		English
LDT	Bowers et al. (1998) Experiment 5	Word (High F, Homophone)	50	3~6	Word (Low F)	500	Non reported	591	608	17	No stats. Significant interaction between relative frequency and prime type	English
LDT	Bowers et al. (1998) Experiment 5	Word (Low F, Homophone)	50	3~6	Word (High F)	500	Non reported	574	570	-4		English
Go/ no go	Davis et al. (1998) Experiment 1	Word (Homophone)	57	4~5	Word	710	Non reported	564	564	0	Adult participants. PHON primes vs. ORTH control primes	English
Go/ no go	Davis et al. (1998) Experiment 1	Nonword (PSH)	57	4~5	Word	710	Non reported	577	560	-17		English
Go/ no go	Davis et al. (1998) Experiment 2	Word (Homophone)	57	4~5	Word	710	Non reported	698	699	1	Child participants. PHON primes vs. ORTH control primes	English
Go/ no go	Davis et al. (1998) Experiment 2	Nonword (PSH)	57	4~5	Word	710	Non reported	746	741	-5		English
LDT	Davis et al. (1998) Experiment 3	Nonword (PSH)	57	4~5	Word	710	Non reported	541	545	4	Adult participants. Prime-target pairs differed in one letter.	English
LDT	Davis et al. (1998) Experiment 3	Nonword (ORTH)	57	4~5	Word	710	Non reported	546	545	-1		English
LDT	Davis et al. (1998) Experiment 3	Nonword (PSH)	57	4~5	Word	710	Non reported	527	540	13	Adult participants. Prime-target pairs differed in two letters.	English
LDT	Davis et al. (1998) Experiment 3	Nonword (ORTH)	57	4~5	Word	710	Non reported	550	540	-10		English
LDT	Davis et al. (1998) Experiment 4	Nonword (PSH)	57	4~5	Word	710	Non reported	768	819	51	Child participants. Prime-target pairs differed in one letter. No stats.	English
LDT	Davis et al. (1998) Experiment 4	Nonword (ORTH)	57	4~5	Word	710	Non reported	769	819	50		English
LDT	Davis et al. (1998) Experiment 4	Nonword (PSH)	57	4~5	Word	710	Non reported	773	820	47	Child participants. Prime-target pairs differed in two letters. No stats.	English
LDT	Davis et al. (1998) Experiment 4	Nonword (ORTH)	57	4~5	Word	710	Non reported	780	820	40		English

Task	**Study**	**Prime lexicality**	**Prime duration**	**Prime length**	**Target**	**Target duration**	**Target N density**	**RT: Related**	**RT: Unrelated**	**Unrelated - related**	**Special manipulations**	**Language**
Naming	Shen & Forster (1999) Experiment 1	Word (Homophone)	50	1	Word	500	Non reported	634	664	30		Chinese
Naming	Shen & Forster (1999) Experiment 1	Word (ORTH)	50	1	Word	500	Non reported	624	664	40*		Chinese
Naming	Shen & Forster (1999) Experiment 1	Word (Homophone)	50	1	Word	500	Non reported	714	731	17		Chinese
Naming	Shen & Forster (1999) Experiment 1	Word (ORTH)	50	1	Word	500	Non reported	699	731	32		Chinese
Naming	Shen & Forster (1999) Experiment 1	Word (ORTH + Homophone)	50	1	Word	500	Non reported	696	731	35		Chinese
LDT	Shen & Forster (1999) Experiment 2	Word (Homophone)	50	1	Word	500	Non reported	578	576	-2		Chinese
LDT	Shen & Forster (1999) Experiment 2	Word (ORTH)	50	1	Word	500	Non reported	541	576	35*		Chinese
LDT	Shen & Forster (1999) Experiment 2	Word (Homophone)	50	1	Word	500	Non reported	537	541	4		Chinese
LDT	Shen & Forster (1999) Experiment 2	Word (ORTH)	50	1	Word	500	Non reported	521	541	20		Chinese
LDT	Shen & Forster (1999) Experiment 2	Word (ORTH + Homophone)	50	1	Word	500	Non reported	517	541	24*		Chinese
LDT	Rastle & Brysbaer (2006) Experiment 1	Nonword (PSH)	58	3~6	Word	Until being responded	8.39	603	616	13*	PSH primes vs. ORTH primes	English
LDT	Rastle & Brysbaer (2006) Experiment 2	Nonword (PSH)	58	3~6	Word	Until being responded	8.39	634	643	9	PSH primes vs. ORTH primes. With PSH nonword foils	English
LDT	Kinoshita & Norris (2012a)	Nonword (PSH)	40	4	Word	2000	9.6	589	599	10*	Target with many Ns	English
LDT	Kinoshita & Norris (2012a)	Nonword (ORTH)	40	4	Word	2000	9.6	595	599	4		English
LDT	Kinoshita & Norris (2012a)	Nonword (PSH)	40	4	Nonword	2000	Non reported	690	693	3		English
LDT	Kinoshita & Norris (2012a)	Nonword (ORTH)	40	4	Nonword	2000	Non reported	674	693	19		English
LDT	Kinoshita & Norris (2012a)	Nonword (PSH)	40	4	Word	2000	0	554	601	47*	Target with few Ns, Significant facilitation from PSH primes over ORTH primes	English
LDT	Kinoshita & Norris (2012a)	Nonword (ORTH)	40	4	Word	2000	0	571	601	30*		English
LDT	Kinoshita & Norris (2012a)	Nonword (PSH)	40	4	Nonword	2000	Non reported	633	627	-6		English
LDT	Kinoshita & Norris (2012a)	Nonword (ORTH)	40	4	Nonword	2000	Non reported	627	627	0		English

PSH: Pseudohomophone, ORTH: Orthographic, PHON: Phonological, F: Frequency, No stats.: Reporting main effects only without statistical analysis for the form priming effect under each condition, *: $p < .05$.

Table 6

The Results of Cross-language Phonological Priming Experiments

Task	Study	Prime condition	Prime duration	Target	Target duration	Target length	Target N	RT: Related	RT: Unrelated	Unrelated - related	Special manipulations	Language
LDT	Gollan et al. (1997) Experiment 1	L1 word (Cognate)	50	L2 word	500	Not reported	Not reported	642	695	53	A significant main effect of prime type, A non- significant interaction between prime type and prime condition	L1 Hebrew, L2 English
LDT	Gollan et al. (1997) Experiment 1	L1 word (Translation)	50	L2 word	500	Not reported	Not reported	676	712	36		
LDT	Gollan et al. (1997) Experiment 2	L1 word (Cognate)	50	L2 word	500	Not reported	Not reported	863	1005	142*	A significant main effect of prime type, a S interaction between prime type and prime condition	L1 English, L2 Hebrew
LDT	Gollan et al. (1997) Experiment 2	L1 word (Translation)	50	L2 word	500	Not reported	Not reported	927	979	52		
LDT	Gollan et al. (1997) Experiment 3	L2 word (Cognate)	50	L1 word	500	Not reported	Not reported	583	592	9	A significant main effect of prime type	L1 Hebrew, L2 English
LDT	Gollan et al. (1997) Experiment 3	L2 word (Translation)	50	L1 word	500	Not reported	Not reported	565	574	9		
LDT	Gollan et al. (1997) Experiment 4	L2 word (Cognate)	50	L1 word	500	Not reported	Not reported	580	584	4		L1 English, L2 Hebrew
LDT	Gollan et al. (1997) Experiment 4	L2 word (Translation)	50	L1 word	500	Not reported	Not reported	586	582	-4		
LDT	Kim & Davis (2003) Experiment 1	L1 word (Cognate)	50	L2 word	500	3.7	11	594	628	34*		L1 Korea, L2 English
LDT	Kim & Davis (2003) Experiment 1	L1 word (Noncognate)	50	L2 word	500	4.8	6	634	674	40*		
LDT	Kim & Davis (2003) Experiment 1	L1 word (Homophone)	50	L2 word	500	3.5	9	634	652	18		
Naming	Kim & Davis (2003) Experiment 2	L1 word (Cognate)	50	L2 word	500	3.7	11	572	600	28		
Naming	Kim & Davis (2003) Experiment 2	L1 word (Noncognate)	50	L2 word	500	4.8	6	664	672	8		
Naming	Kim & Davis (2003) Experiment 2	L1 word (Homophone)	50	L2 word	500	3.5	9	597	617	20*		
Semantic	Kim & Davis (2003) Experiment 4	L1 word (Cognate)	50	L2 word	500	3.7	11	724	776	52*		
Semantic	Kim & Davis (2003) Experiment 4	L1 word (Noncognate)	50	L2 word	500	4.8	6	728	786	58*		
Semantic	Kim & Davis (2003) Experiment 4	L1 word (Homophone)	50	L2 word	500	3.5	9	775	774	-1		
Naming	Zhou et al. (2010) Experiment 1	L1 word (Homophone)	50	L2 word	Until being responded	2~5	Not reported	621	631	10	A significant main effect of prime type, A non-significant interaction between prime type and proficiency	L1 Chinese, L2 English high prof.
Naming	Zhou et al. (2010) Experiment 1	L1 word (Homophone)	50	L2 word	Until being responded	2~5	Not reported	660	674	14		L1 Chinese, L2 English low prof.
Naming	Zhou et al. (2010) Experiment 2	L2 word (Homophone)	50	L1 word	Until being responded	7.20 strokes	Not reported	544	551	7	A significant main effect of prime type, A non-significant interaction between prime type and proficiency	L1 Chinese, L2 English high prof.
Naming	Zhou et al. (2010) Experiment 2	L2 word (Homophone)	50	L1 word	Until being responded	7.20 strokes	Not reported	585	594	9		L1 Chinese, L2 English low prof.
LDT	Zhou et al. (2010) Experiment 3	L1 word (Homophone)	50	L2 word	Until being responded	2~5	Not reported	652	666	14	A significant main effect of prime type, A non-significant interaction between prime type and proficiency	L1 Chinese, L2 English high prof.
LDT	Zhou et al. (2010) Experiment 3	L1 word (Homophone)	50	L2 word	Until being responded	2~5	Not reported	631	660	29		L1 Chinese, L2 English low prof.
LDT	Zhou et al. (2010) Experiment 4	L1 word (Homophone)	50	L1 word	Until being responded	7.20 strokes	Not reported	574	599	25	A significant main effect of prime type, A non-significant interaction between prime type and proficiency	L1 Chinese, L2 English high prof.
LDT	Zhou et al. (2010) Experiment 4	L1 word (Homophone)	50	L2 word	Until being responded	7.20 strokes	Not reported	534	544	10		L1 Chinese, L2 English low prof.

Task	Study	Prime condition	Prime duration	Target	Target duration	Target length	Target N density	RT: Related	RT: Unrelated	Unrelated - related	Special manipulations	Language
LDT	Nakayama et al. (2012)	L1 word (Cognate)	50	L2 word (Low F)	Until being responded	4.6	6.4	608	701	93	No stats. Cognate priming was more sensitive to proficiency and target frequency than homophone priming.	L1 Japanese, L2 English high prof.
LDT	Nakayama et al. (2012)	L1 word (Homophone)	50	L2 word (Low F)	Until being responded	4.6	6.4	667	701	34		
LDT	Nakayama et al. (2012)	L1 word (Cognate)	50	L2 word (High F)	Until being responded	4.7	6.4	577	633	56		
LDT	Nakayama et al. (2012)	L1 word (Homophone)	50	L2 word (High F)	Until being responded	4.7	6.4	622	633	11		
LDT	Nakayama et al. (2012)	L1 word (Cognate)	50	L2 word (Low F)	Until being responded	4.6	6.4	651	774	123		L1 Japanese, L2 English low prof.
LDT	Nakayama et al. (2012)	L1 word (Homophone)	50	L2 word (Low F)	Until being responded	4.6	6.4	741	774	33		
LDT	Nakayama et al. (2012)	L1 word (Cognate)	50	L2 word (High F)	Until being responded	4.7	6.4	595	704	109		
LDT	Nakayama et al. (2012)	L1 word (Homophone)	50	L2 word (High F)	Until being responded	4.7	6.4	661	704	43		
LDT	Nakayama et al. (2013) Experiment 1	L1 word (Cognate)	50	L2 word (Low F)	Until being responded	5.1	Not reported	614	694	80	The cognate priming advantage over the priming by noncognate translation primes was not modulated by target frequency and proficiency.	L1 Japanese, L2 English high prof.
LDT	Nakayama et al. (2013) Experiment 1	L1 word (Translation)	50	L2 word (Low F)	Until being responded	5.1	Not reported	628	678	50		
LDT	Nakayama et al. (2013) Experiment 1	L1 word (Cognate)	50	L2 word (High F)	Until being responded	5.2	Not reported	576	610	34		
LDT	Nakayama et al. (2013) Experiment 1	L1 word (Translation)	50	L2 word (High F)	Until being responded	5.1	Not reported	584	615	31		
LDT	Nakayama et al. (2013) Experiment 1	L1 word (Cognate)	50	L2 word (Low F)	Until being responded	5.1	Not reported	625	731	106		L1 Japanese, L2 English low prof.
LDT	Nakayama et al. (2013) Experiment 1	L1 word (Translation)	50	L2 word (Low F)	Until being responded	5.1	Not reported	643	726	83		
LDT	Nakayama et al. (2013) Experiment 1	L1 word (Cognate)	50	L2 word (High F)	Until being responded	5.2	Not reported	568	655	87		
LDT	Nakayama et al. (2013) Experiment 1	L1 word (Translation)	50	L2 word (High F)	Until being responded	5.1	Not reported	604	660	56		
LDT	Nakayama et al. (2013) Experiment 2	L2 word (Cognate)	50	L1 word	Until being responded	3.5 characters	Not reported	513	543	30*	A significant interaction between prime type and proficiency	L1 Japanese, L2 English high prof.
LDT	Nakayama et al. (2013) Experiment 2	L2 word (Cognate)	50	L1 word	Until being responded	3.5 characters	Not reported	540	555	15*		L1 Japanese, L2 English low prof.
LDT	Ando et al. (2014) Experiment 1	L1 (Homophone)	50	L2 word	1500	2~7	Not reported	711	728	17*		L1 Japanese, L2 English
LDT	Ando et al. (2014) Experiment 2	L1 (Homophone)	50	L2 word (High F)	2500	2~6	Not reported	563	585	22	A significant main effect of prime type, a significant interaction between prime type and target F	L1 Japanese, L2 English
LDT	Ando et al. (2014) Experiment 2	L1 (Homophone)	50	L2 word (Low F)	2500	2~6	Not reported	631	644	13		L1 Japanese, L2 English
LDT	Ando et al. (2015)	L1 (Homophone)	50	L2 word (High F)	1500	4.6	6.4	852	872	20	A significant main effect of prime type, a significant interaction between prime type and target F	L1 Japanese, L2 English
LDT	Ando et al. (2015)	L1 (Homophone)	50	L2 word (Low F)	1500	4.7	6.4	765	798	33		L1 Japanese, L2 English

F: Frequency, No stats.: Reporting main effects only without statistical analysis for the form priming effect under each condition, prof.: proficiency, *: $p < .05$.

2.2.3. Cross-language Masked Phonological Priming

As noted in an earlier section, Bijeljac-Babic et al. (1997) showed that the L2 orthographic neighbor word prime could produce inhibition for the L1 word target if French-English bilingual participants' proficiency was high. The authors interpreted this finding based on the Bilingual Interactive Activation model (Van Heuven, Dijkstra, & Grainger, 1998). This model supposes that the mechanism of interactions between sub-lexical and lexical units that is posited in the Interactive Activation model (McClelland & Rumelhart, 1981) is not limited to within a language but works across bilinguals' two different languages. Thus, the form inhibition found in their study was considered to have come from simultaneous activation of orthographic representations of neighbors in both languages.

This cross-language activation was similarly observed in the domain of phonological representations. Duyck (2005), for example, showed that in an LDT, Dutch-English bilinguals' responses to the L2 word target (e.g., *back*) were facilitated when masked pseudo-homophone primes (e.g., a Dutch pseudo-homophone *ruch*) whose pronunciation was the same as the L1 translation equivalent of the target (the Dutch word *rug* meaning "back") were presented. This simultaneous activation of cross-language phonology-semantic networks is surprising, but interestingly, in all eight experiments in Duyck's study (2005), the pronunciation of the prime always affected target identification through the mediation of semantics. This is because cross-language (pseudo) homophones are likely to have similar orthographic forms to the target if bilinguals' two languages use the same script, and Duyck (2005) wanted to avoid the influence of orthographic overlap between the prime and target. The necessity for controlling this factor is particularly greater in the LDT since lexical competition is supposed to be stronger when the response task suggests the importance of orthographic form information, which is the case for the LDT (Kim & Davis, 2003).

However, if bilinguals' two languages use different scripts, there would be less interference of the orthography (i.e., facilitation from letter overlap or inhibition from lexical competition between within- and cross-language neighbors) because the prime and the target do not have similar orthographic forms. Therefore, it is possible to assume that it would be easier to observe the contribution of phonology in phonological priming experiments if participants are different-script bilinguals. Supporting this hypothesis, previous studies have observed robust facilitative phonological priming across two languages that do not share orthographies.

The results of previous masked, different-script, cross-language phonological priming studies are summarized in Table 6. Even though Kim and Davis (2003) reported that the 18 ms facilitative cross-language homophone priming observed in their study was non-significant, subsequent studies (Ando, Jared, Nakayama, & Hino, 2014; Ando, Matsuki, Sheridan, & Jared, 2015; Nakayama, Sears, Hino, & Lupker, 2012; Zhou, Chen, Yang, & Dunlap, 2010) discovered that a masked L1 word prime with pronunciation similar to the L2 word target yielded significant facilitation. Studies on cognates, cross-language word pairs that have both similar pronunciations and meanings (i.e., loan words), that posited that cognate priming is a combination of phonological and semantic priming (Gollan, Forster, & Frost, 1997; Nakayama et al., 2012; Nakayama, Sears, Hino, & Lupker, 2013) also found that cognate priming was always stronger than noncognate cross-language translation priming. This finding also supports the idea that a phonological signal extracted from the prime facilitates target identification if the prime and target have similar phonological forms. In addition, the same research also documents that this phonological priming occurred even in the L2-L1 direction (i.e., when the mask is an L2 homophone for a L1 target) (Nakayama et al., 2013; Zhou et al., 2010), and for logographic languages (Chinese and Japanese Kanji) in which each character has a meaning and phonological assembly using sub-lexical units (i.e., letters) is impossible (Ando et al., 2014; Nakayama et al., 2013; Zhou et al., 2010).

In terms of the locus of phonological priming, even though Zhou et al. (2010) thought that the cross-language homophone priming found in their study was lexical because there is no grapheme-phoneme conversion rule in the language that they used (Chinese), subsequent studies (Ando et al., 2014; Nakayama et al., 2012, 2013) claimed that different-script, cross-language phonological priming, including that involving logographic languages, is prelexical (i.e., caused by sub-lexical phonology), as in same-script languages (e.g., Van Heuven et al., 1998). The evidence for the latter claim was found in the discovery that phonological priming was not significantly affected by L2 proficiency or frequency of L2 word stimuli unlike cross-language translation (semantic) priming. This argument was further supported by a neurological study (Ando et al., 2015) that used Event Related Potential (ERP) data, the data collected while tracking changes in voltage generated in the brain structures in response to specific events or stimuli over time. That study reports that the effect of phonological priming was observed before a target frequency effect emerged in the ERP data that the researchers collected during participants' lexical decisions.

In brief, more robust facilitative phonological priming has been found for different-script cross-language homophone pairs than for within-language L1 homophone pairs. Of course, cross-language phonological priming is not identical to L2-L2 phonological priming because there can be stronger lexical competition if both the prime and the target are from the same language. However, as reviewed above, form inhibition has been found to be weak between L2 words. Therefore, it is reasonable to predict that because the interference of cross-language orthographic neighbors will be relatively weaker if bilinguals' two languages use different scripts, within-language L2 phonological priming would be more readily observable in different-script L2 learners. This suggests that masked priming experiments with prime-target pairs with difficult phonological contrasts offer a useful tool for exploring how these L2 words are differently processed from other L2 words without a confusing contrast. Furthermore, in combination with a translation and familiarity rating task, the findings from this study may help us tease apart the effect of fuzzy phonological representations from that of fuzzy orthographic representations on form facilitation. If a participant feels that he/she is familiar with the prime (e.g., *read*) but wrongly provides the L1 translation equivalent of another word (e.g., *lead*) in the translation task, it will indicate that the form facilitation observed in a form priming experiment might have resulted from confusion of the orthographic form of the prime with that of the target. However, if it turns out that this participant can correctly identify the prime and the target in the translation task but still shows form facilitation for this prime-target pair, it may suggest that the indistinguishable phonological forms of the prime and the target that are registered in L2 learners' memory are the origin of form facilitation. In other words, in the latter case, the form facilitation will mean that the masked phonological priming observed in previous research between within- and between-language homophones might have arisen even for L2 near-homophones. Based on this reasoning, the present study examines whether prime-target pairs with confusing L2 phonological contrasts would produce different patterns of priming from other pairs without a difficult contrast.

Chapter 3. Methodology

3.1. The Present Study

The FLR hypothesis (Bordag et al., 2021; Gor et al, 2021) posits that due to the fact that the L2 lexicon is acquired after the L1 lexicon, phonological and orthographic L2 word forms are not efficiently encoded. Thus, these fuzzy forms are predicted to facilitate access to their orthographic and phonological neighbors through sublexical pre-activation, while they are weakly engaged in lexical competition. Although the FLR hypothesis predicts that the quality of form representations improves along with L2 experience, successful encoding of L2 word forms with difficult phonological contrasts may be more challenging than other words without a confusing contrast. Based on this understanding of the literature, the present study hypothesized that the development of the quality of L2 form representations with and without a nonnative contrast would be observable using the masked orthographic neighbor priming method. Specifically, the current study predicted that (a) at an initial stage of L2 lexical development (i.e., in low-proficiency L2 speakers), form facilitation in an orthographic neighbor priming experiment would take place due to underdeveloped orthographic representations. However, as L2 experience increase, orthographic representations were expected to become robust enough to produce at least a limited level of form inhibition, as observed in the null form priming effect where sublexical facilitation and lexical inhibition cancel each other out. Nevertheless, the establishment of fully specified form representations for words with nonnative phonological contrasts was expected to require more L2 experience. Therefore, (b) medium-proficiency L2 learners, who would show some form inhibition for prime-target pairs without a difficult phonological contrast (e.g., *dear-TEAR*), were predicted to still show form facilitation for other prime-target pairs with a confusing phonological contrast (e.g., *read-LEAD*). This phenomenon was expected to occur either because of fuzzy orthographic encodings of L2 words with nonnative phonological contrasts that are caused by fuzzy phonological representations, or because of the indistinguishable phonological forms of minimal pairs with difficult phonological contrasts that are both activated during word reading. Finally, since the FLR hypothesis posits that the lexical quality of L2 words improves and eventually becomes functionally comparable to L1 words, this study predicted that (c) high-proficiency L2 learners would show form inhibition for L2 word prime-target pairs both with and without nonnative phonological contrasts.

In order to test these three hypotheses, five prime conditions were created for a masked-primed LDT as illustrated in Table 7 below. The amount of orthographic overlap, as well as overlap position, was the same in the first (e.g., *read-LEAD*), second (e.g., *full-Pull*) and third (e.g., *dear-TEAR*) conditions. The first two conditions consisted of prime-target pairs with nonnative phonological contrasts. On the other hand, the third condition was composed of pairs without difficult L2 phonological contrasts.

The fourth condition (*week-WEAK*) was created to compare the pattern of priming observed under the homophone condition with those patterns detected under the /l/-/ɹ/ and /f/-/p/ contrast conditions. In previous research, the facilitative within-language homophone priming effect has usually been thought to come from the pre-activation of the phonemes of the target by the prime. However, the facilitation for L2 homophone pairs may rather be due to fuzzy orthographic representations caused by the interference of the same sound for two different words in the development of orthographic representations (see Section 2.2.3). Even though it would be difficult to determine whether the form priming effect for homophone pairs comes from phonology or orthography, it was expected that if the pattern of priming found under the /f/-/p/ and /l/-/ɹ/ conditions was similar to that observed under the homophone condition, at least it might indicate that the form priming for minimal pairs with a nonnative contrast and that for homophone pairs show the effects of fuzzy form representations.

The fifth condition (*play-PLAN*) was included in preparation for an unexpected situation. To be more specific, if, contrary to our prediction, even L1 speakers did not show significant form inhibition under the first three conditions (e.g., *read-LEAD, full-PULL* and *dear-TEAR*) in which the prime and the target differed only in the first letter, the reason for the weak lexical competition under these conditions could be inferred based on the pattern and strength of an orthographic priming effect observed under the fifth condition. In the fifth condition, all prime-target pairs shared the same first letter and differed in the middle letter positions or the last letter position.

This section presents the reasons for including each of these five prime conditions and for using other tasks in addition to an LDT. The next section will then describe the specific characteristics of the stimuli used under each prime condition in the LDT and of the other tasks employed in the current study.

The first prime condition was for prime-target pairs with the /r/-/l/ phonological contrast (e.g., *read-lead*). The /l/-/ɹ/ contrast was chosen because this contrast is not discriminated in Korean (Schmidt, 1996), and it is known to be difficult to learn for late Korean L2 learners of English (Borden, Gerber, & Milsark, 1983; Jamieson & Yu, 1996; Smith, 2001). Since Jamieson and Yu (1996) documented that Korean L2 speakers of English identified the /l/-/ɹ/ contrast more accurately when they were aurally presented at the final singleton position of words (i.e., …verb + r/l as in *steer-steel*), to minimize the unwanted effects of allophonic differences, all prime-target pairs used under this prime condition always differed in the first letter (e.g., *read-LEAD*). This rule was also equally applied to the second and third prime conditions.

Other minimal pairs with another nonnative confusing contrast /f/-/p/ (e.g., *full-PULL*) were used for the second condition. Research on Korean L2 learners' acquisition of this contrast has been scarce, unlike that for the /l/-/ɹ/ contrast. However, Park and De Jong (2008) and Schmidt (1996) report that when late Korean L2 learners were asked to label aurally presented English phones using Korean consonants, they frequently labeled these two distinct English phones as a single Korean consonant ("ㅍ"), indicating that these two L2 consonants are not discriminated in Korean. The reason for including this condition was to examine the generalizability of the impact of difficult phonological contrast on orthographic encoding. Put differently, this study tried to make sure that a stronger form facilitation occurs not only for words with the /l/-/r/ contrast but also for words with other nonnative contrasts (e.g., /f/-/p/). However, there was a small difference between the /l/-/ɹ/ and /f/-/p/ contrasts. That is, whereas Korean participants in Schmidt's study (1996) responded that /l/ and /ɹ/ were perceived as equally distant from a corresponding Korean consonant ("ㄹ"), they felt that /p/ was closer to the corresponding Korean consonant "ㅍ" than /f/. According to L2 sound perception research (Best & Tyler, 2007; Flege, 2003), the former contrast is hypothesized to be more difficult to acquire. Therefore, it was thought to be possible that the phonological forms of words with the /f/-/p/ contrast would be relatively more quickly learned and thus stop producing stronger facilitation at a relatively earlier stage than other minimal pairs with the /r/-/l/ contrasts. However, whether this small difference would indeed affect the rates of nonnative contrast acquisition and produce different results for the /l/-/ɹ/ and /f/-/p/ contrasts in this masked priming experiment remained to be explored.

The third prime condition consisted of minimal pairs without nonnative contrasts (e.g., *dear-TEAR*). Although the initial attempt was made to collect prime-target pairs with only a single phonological contrast (e.g., /d/-/t/), it was impossible to find a sufficient number of those pairs that were expected to be known to most L2 learners. Except for those with the /g/-/k/ contrast, such pairs are rare. However, the /g/-/k/ contrast was not chosen because one-to-one correspondence between a phoneme and a grapheme was not guaranteed for the phoneme /k/ (i.e., the letter "c" can be pronounced as /s/ as in *cease* as well as /k/ as in *clip*). Therefore, it was finally decided that stimuli for this prime condition would contain minimal pairs not only with the /d/-/t/ contrast but also those with the /b/-t/ contrast (e.g., *took-BOOK*). Prime-target pairs used under this condition would represent the improvement in the quality of orthographic representations of L2 words without difficult phonological contrasts. Specifically speaking, under this condition, low-proficiency L2 speakers were expected to show form facilitation. On the other hand, medium- and high-proficiency L2 speakers were expected to show a null priming effect or even form inhibition.

The prime-target pairs for the fourth priming condition were prepared to examine whether masked homophone priming occurs in the L2 (e.g., *week-WEAK*). The amount of overlap in orthographic forms between the prime and the target was not greatly different from that under the other conditions. 13 prime-target pairs differed only in one letter from the stimuli used under the other conditions, and there were only three pairs with less overlap (*here-hear*, *one-WON* and *sight-SITE*). Although inclusion of these three pairs was not ideal, they were selected because there were no more homophone pairs differing in only one letter which were predicted to be known to most L2 participants. The linguistic properties (log frequency, neighborhood density and mean bigram frequency) of the prime and the target used for this condition were also comparable to those used under the other conditions ($p > .10$ in all ANOVAs and pair-wise comparisons) (see Appendix E). A noteworthy difference in stimuli between this and the other conditions was that most stimuli under this condition were less transparent in terms of grapheme-phoneme conversion (e.g., *see-SEA* or *toe-TOW*). It was difficult to predict what influence this difference would exert on priming effects across the different participant groups. However, it was thought that the influence of this difference would not be very great for experienced readers of English such as participants in the current study[10]. Based on this assumption, if facilitative priming would be stronger under this condition compared to the other orthographic priming conditions in which words without nonnative phonological contrasts were used (e.g., *dear-TEAR* or *play-PLAN*), the additive priming observed under this condition would be interpreted to have come from activation of phonology of the masked homophone prime.[11] This result would indicate that L2 participants recruited for the current study were able to activate phonological codes of the masked prime. Thus, findings obtained from this condition were expected to be useful for inferring the source of facilitative priming observed under the first (e.g., *read-LEAD*) and second (e.g., *full-PULL*) conditions. To be more specific, if facilitative priming (which will be compared with the priming for *dear-TEAR* or *play-PLAN*) was to be observed for *see-SEA* but not for *read-LEAD*, this would indicate that the phonological contrast /l/-/ɹ/ that was registered in L2 speakers' memory might be discriminable. On the other hand, if

---

[10] Even though one of the three groups of L2 speakers is labelled as "low-proficiency L2 speaker", this does not mean that these participants were beginner L2 learners. The label was chosen for easier discrimination of three L2 speaker groups in the present study. Low-proficiency L2 speakers who were recruited in universities in Korea had substantial experience of reading English texts because they had received English education for longer than 9 years at primary and secondary schools.

[11] Regarding the position of the difference between the prime and the target, the stimuli used under the homophone priming condition were more similar to those used under the fifth priming condition (e.g., *play-PLAN*) than the other three conditions (e.g., *read-lead*, *full-pull* and *dear-TEAR*). Specifically speaking, all prime-target pairs under the fifth condition differed either in the middle letter positions or the last letter position whereas the pairs under the other three conditions always differ in the first letter. Similarly, out of 13 homophone pairs that differed in only one letter, eight pairs differed in the middle letter positions (e.g., *week-WEAK*), and four pairs differed in the last letter positions (e.g., *pole-POLL*). Only one pair differed in the first letter position (e.g., *sell-CELL*).

facilitative priming was not to be observed for *see-SEA* whereas facilitative priming for *read-LEAD* was to be observed, this would indicate that facilitation for the latter pair came from something other than the activation of phonological information of the masked prime *read*. All the homophone pairs used under the fourth condition were from the Appendixes in Lukatela and Turvey's study (1994).

The final prime condition was included to examine whether there would be a difference in the size of form priming depending on the position of overlap between the prime and the target. To be more specific, under the first three priming conditions of the current study (the conditions for the /l/-/ɹ/, /f/-/p/, and /b/- or /d/-/t/ contrasts), the prime-target pairs always differed in their first-letter position. However, since previous masked form priming studies have not closely examined the effect of overlap positions between the prime and the target, it was considered possible that lexical competition would be weaker if the prime-target pairs differed only in their first letter (as in *read-LEAD*). For instance, Colombo (1986) claimed that readers begin to activate relevant words while reading the initial part of the word. Thus, according to her argument, inhibition would be stronger for prime-target pairs sharing initial letters (e.g., *captain-captive*) because in this case, the prime would be activated fast enough to initiate inhibition before lexical access to the target (see also Marslen-Wilson, 1987 for a similar argument for *early selection* during spoken word identification). In addition, previous unmasked priming studies (Colombo, 1986; Slowiaczek & Hamburger, 1992) indeed obtained findings that supported Colombo's claim. Therefore, if Colombo's assumption was applied to the mechanism of the masked priming paradigm, lexical competition between orthographic neighbor pairs that differ in the word initial and share rhymes (e.g., *dear-TEAR*) could be weak even in L1 speakers. In other words, the strong inhibition for orthographic neighbor pairs reported in previous studies (see Section 2.1.1) could have mostly been driven by pairs differing in their middle letter positions or final letter position. If so, this would not be good for the present study in which the three critical prime conditions always differed only in the first letter position. This is because even prime-target pairs without a nonnative contrast would not show strong inhibition regardless of L2 participants' proficiency, nor would they be significantly differentiated from pairs with a confusing contrast.

To the best of our knowledge, only one study has addressed a similar issue arising here. Frisson, Bélanger and Rayner (2014, Experiment 2) compared the size of masked priming effects under three different conditions: High orthographic and high phonological overlap (e.g., *track–crack*), high orthographic and low phonological overlap (e.g., *bear–gear* or *awash-abash*) and low orthographic and high phonological overlap (e.g., *fruit–chute* or *fellow-ghetto*). They then further divided the first condition into two: high orthographic and high phonological end overlap (P+O+ rhyme, e.g., *track–crack*) and high orthographic and high phonological beginning overlap (P+O+ unrhyme, e.g., *swoop-swoon*). The results of their experiment are presented in Table 8. We are interested in the comparison between the P+O+ rhyme and P+O+ unrhyme conditions. Frisson and his colleagues (2014) reported that form inhibition under the P+O+ rhyme condition was significant but not under the P+O+ unrhyme condition. These findings indicate that in the present study, significant inhibition is likely to be observed for *down-TOWN* if participants have robust orthographic representations. However, since there is only one study, and since it is uncertain whether the same findings are observable in L2 speakers, the present study included a prime condition in which prime-target pairs without a confusing phonological contrast differed in a median or the last letter (e.g., *play-PLAN*) in order to examine how the strength of orthographic priming differs depending on overlap position.

Table 8

The results of Frisson, Bélanger and Rayner' Experiment 2 (2014)

	P+O+ (rhyme)		P+O+ (nonrhyme)		P-O+		P+O-	
	Control (*luigh-crack*)	Related (*track-crack*)	Control (*blydt-swoon*)	Related (*swoop-swoon*)	Control (*qulk-gear*)	Related (*bear-gear*)	Control (*neelk-chute*)	Related (*fruit-chute*)
RTs	624	642	625	621	636	652	644	653
Priming	-18*		4		-16*		-9	

- P+O+ (rhyme): High orthographic and high phonological beginning overlap
- P+O+ (nonrhyme): High orthographic and high phonological end overlap
- P-O+: High orthographic and low phonological overlap
- P+O-: Low orthographic and high phonological overlap

The use of five priming conditions addressed above was beneficial for teasing apart the effect of representations from that of processing speed. Previous research found that different properties of the prime are activated according to different intervals between the prime and the target, such that different patterns of priming effects can be observed from the same prime depending on distinct SOAs. For instance, it was found that the visual form prime produced facilitation (Ferrand & Grainger, 1992, 1993; Perfetti & Tan, 1998) with shorter SOAs (e.g., 33 ~ 50 ms) but gave way to inhibition with longer SOAs (e.g., 67 ms). On the other hand, the phonological prime was found to produce facilitation with relatively longer SOAs (e.g., 67 ms). Moreover, Gollan and her collaborators (1997) suggested that different patterns of priming from the L1 and L2 primes could be because of the difference in the speed of processing the prime. According to these previous findings and suggested speculations, any differences in patterns of priming between L1 and L2 speakers could be considered a result of the relative processing speed in the two languages. However, through employing the different priming conditions, it was possible to control the influence of processing speed because all stimuli that had similar linguistic properties across the conditions were presented with the same SOA. Thus, if any differences in the patterns of priming across the five priming conditions were to be observed, these differences were expected to be attributed to the effects of lexical representations.

In addition to the LDT, the quality of orthographic encoding and vocabulary size were measured using spelling dictation and spelling recognition tasks (Andrews et al., 2020) and a vocabulary test (Shipley, 1940). If the reason for the different patterns of orthographic neighbor priming observed across three L2 speaker groups (i.e., low-, medium- and high-proficiency L2 speakers) was due to the different levels of precision in their orthographic representations or different vocabulary sizes, it was expected that form priming would be modulated by either orthographic precision or vocabulary size.

In addition, a word identification task and a phoneme identification task were used in order to examine the relationship between sound perception and phonological priming for near-homophone prime-target pairs. In the word identification task, participants heard each word that was used as prime or target under the three critical priming conditions (the conditions for the /l/-/ɹ/, /f/-/p/, and /d/- or /b/-/t/ contrasts) of the LDT (e.g., *read*).

Finally, a cloze test (Brown, 1980) was used as a measure of global proficiency. There were two reasons why this proficiency test was used. First, even though TOEIC scores were used for screening of low- and medium-proficiency L2 speakers during recruitment (see below for details of the recruiting procedure), high-proficiency participants who had lived in the U.S. did not have a score on this standardized English proficiency test. Therefore, we needed an objective proficiency measure to show that the three groups of recruited L2 participants indeed had different levels of proficiency. Second, although it was required for low- and medium-proficiency participants to have a TOEIC score that had been obtained within two years, their scores might not represent the level of their proficiency at the moment of the experiment, particularly if they took a TOEIC long before the experiment. For that reason, we expected that by employing a cloze test, we would be able to accurately measure L2 speakers' proficiency when they took part in the current study.

In conclusion, the present study was driven by the following four research questions.

1. For L2 speakers, what problems in sound identification of L2 nonnative phonological contrasts are revealed by word and phoneme identification tasks?
2. Do three groups of L2 speakers with different levels of proficiency show different patterns of orthographic form priming?
3. Is a relatively more delayed development of form representations of words with difficult nonnative contrasts (compared to that of other words without a difficult nonnative contrast) observable in an orthographic masked priming experiment?
4. What aspects of proficiency (i.e., orthographic precision, vocabulary size or accurate sound perception) modulate the patterns or strengths of orthographic form priming?

Table 7

Characteristics of the Stimuli for the LDT

Word targets (*n* = 80)										Nonword targets (*n* = 80)	
Homophone (*n* = 16)		/l/ - /ɹ/ (*n* = 16)		/p/ - /f/ (*n* = 16)		/d/ - /t/ or /b/ - /t/ (*took – BOOK)* (*n* = 16)		Orthographic neighbors at do not differ in the first letter (*n* = 16)			
○ Length: 3.8 ○ Log F: 3.13 ○ ORTH N: 12.4 ○ Mean bigram F: 3,381		○ Length: 3.8 ○ Log F: 2.98 ○ ORTH N: 16.3 ○ Mean bigram F: 3,768		○ Length: 3.9 ○ Log F: 2.89 ○ ORTH N: 14.1 ○ Mean bigram F: 3,836		○ Length: 3.8 ○ Log F: 2.99 ○ ORTH N: 15.3 ○ Mean bigram F: 3,486		○ Length: 3.9 ○ Log F: 3.06 ○ ORTH N: 15.6 ○ Mean bigram F: 3,960		○ Length: 3.9 ○ ORTH N: 6.6	
Word primes (40 related and 40 unrelated primes in each list)										Word primes (40 related and 40 unrelated primes)	
Related (*week-WEAK*)	Unrelated (*hurt-WEAK*)	Related (*read-LEAD*)	Unrelated (*wish-LEAD*)	Related (*full-PULL*)	Unrelated (*sick-PULL*)	Related *(dear-TEAR)*	Unrelated *(boys-TEAR)*	Related (*play-PLAN*)	Unrelated (*best-PLAN*)	Related (*name-MAME*)	Unrelated (*long-MAME*)
○ Length: 3.8 ○ Log F: 3.85 ○ ORTH N: 13.9 ○ Mean bigram F: 3,693	○ Length: 3.8 ○ Log F: 3.81 ○ ORTH N: 13.8 ○ Mean bigram F: 3,778	○ Length: 3.8 ○ Log F: 3.61 ○ ORTH N: 15.9 ○ Mean bigram F: 3,840	○ Length: 3.8 ○ Log F: 3.63 ○ ORTH N: 14.4 ○ Mean bigram F: 3,896	○ Length: 3.9 ○ Log F: 3.66 ○ ORTH N: 13.1 ○ Mean bigram F: 3,622	○ Length: 3.9 ○ Log F: 3.66 ○ ORTH N: 13.2 ○ Mean bigram F: 3,607	○ Length: 3.8 ○ Log F: 3.87 ○ ORTH N: 14.9 ○ Mean bigram F: 3,681	○ Length: 3.8 ○ Log F: 3.87 ○ ORTH N: 14.6 ○ Mean Bigram F: 3,794	○ Length: 3.9 ○ Log F: 3.95 ○ ORTH N: 15.3 ○ Mean bigram F: 3,541	○ Length: 3.9 ○ Log F: 3.94 ○ ORTH N: 13.8 ○ Mean bigram F: 3,227	○ Length: 3.8 ○ Log F: 3.78 ○ ORTH N: 11.1 ○ Mean bigram F: 3,157	○ Length: 3.8 ○ Log F: 3.78 ○ ORTH N: 11.8 ○ Mean bigram F: 3,178

Log F: Log SUBTLEX-US Frequency, ORTH N: Number of orthographic neighbors (neighborhood density), Mean bigram F: Mean bigram frequency

3.2. Participants

One group of native English (L1) speakers and three groups of Korean L2 learners of English (L2 speakers) with different levels of proficiency were recruited through a flier posted on Social Network Services. 30 native English speakers were recruited in the United States to serve as the control group in the present study. Three groups of Korean L2 learners of English were recruited, all required to have no experience studying in an English-only institution for longer than three months before age 18. This was because the way early L2 learners represent L2 words in an L2 immersion environment is assumed to be different from how they represent L2 words in a formal learning environment, and different as well from late L2 learners' representations in immersion contexts (e.g., Bordag et al., 2021; Jiang, 2000; Jiang & Pae, 2021). The present study explores the nature of the L2 lexical representations that are established in a formal learning environment or by late L2 learners.

Both low- and medium-proficiency L2 participants were recruited from 10 different universities in Korea. Medium-proficiency participants (n = 30) were required to have obtained a TOEIC score higher than 850 within two years, and low-proficiency participants (n = 30) were required to have obtained a TOEIC score between 650~850. The lower cut-off score for the latter group (650 on TOEIC) was to make sure that even low-proficiency participants knew most of the stimuli that were used in the masked priming LDT. High-proficiency participants (n = 30) were required to have lived in the U.S. for longer than 5 years after age 18 while regularly using English for their work or study. They were also required to have experience studying in the U.S. to obtain an academic degree higher than a bachelor's degree. This requirement was set to recruit participants with sufficient reading experience. These participants had lived in the U.S. for 10.7 years on average.

Table 9

Participants' Biographical Information

	Age	AOA	LOR	Education	Proficiency
Low-prof. L2 speakers (Female: 23, male: 7)	24.0 (18 ~ 39)	8.0 years (5 ~ 12 years)	-	15.2 years (12 ~ 18 years)	23.3 (15 ~ 33)
Mid-prof. L2 speakers (Female: 28, male: 2)	25.5 (21 ~ 39)	7.8 years (4 ~ 11 years)	-	15.5 years (12 ~ 20 years)	31.4 (14 ~ 43)
High-prof. L2 speakers (Female: 22, male: 8)	35.0 (26 ~ 49)	9.2 years (2 ~ 12 years)	10.7 years (5 ~ 23 years)	19.5 years (15 ~ 26 years)	36.9 (13 ~ 48)
L1 speakers (Female: 22, male: 8)	23.9 (18 ~ 49)	-	-	16.0 years (12 ~ 21 years)	-

Participants' biographical information is presented in Table 9, and the results of statistical tests for group comparisons in consideration of biographical information are illustrated in Appendix E. In brief, low- and mid-proficiency L2 speakers and L1 speakers were all comparable to each other in terms of age and length of formal education. However, high-proficiency participants were significantly older and had received more formal education than the other groups. Ages of acquisition (AOA) of low- and mid-proficiency L2 speakers were not significantly different. On the other hand, high-proficiency L2 speakers began to learn English in Korea at significantly older ages than the other two L2-speaker groups. An additional one-way ANOVA test on the cloze test scores revealed a significant difference in proficiency (measured with a cloze test) across the three L2 speaker groups, $F(2, 87) = 27.64$, $p < .001$ (see Figure 1). For post-hoc group comparisons, Bonferroni-adjusted statistical tests were conducted. The results indicated that the differences in proficiency between low- and mid-proficiency L2 speakers ($p < .001$), between low- and high-proficiency L2 speakers ($p < .001$) and between mid- and high-proficiency L2 speakers ($p = .011$) were all significant.

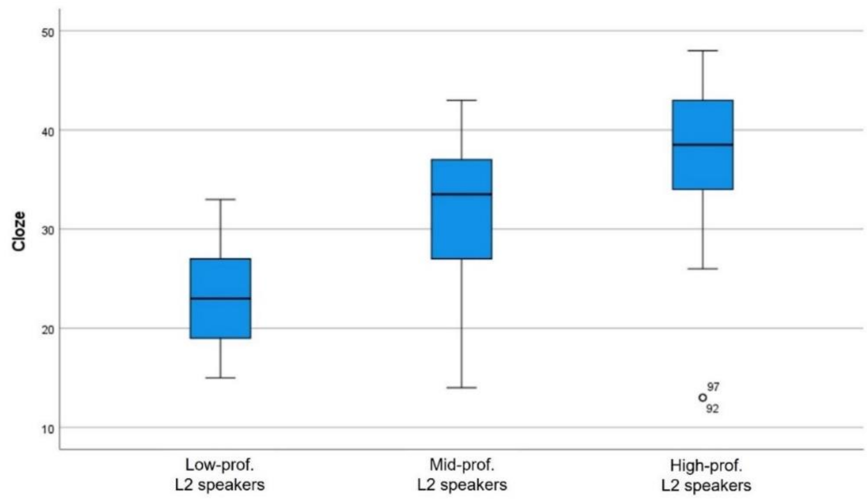

Figure 1. Proficiency (Cloze Test Scores) of Three L2 Speaker Group.

3.3. Materials

3.3.1. Lexical Decision Task

The linguistic characteristics of the stimuli that were obtained from the "complete lexicon" of the English Lexicon Project (Balota et al., 2007) are illustrated in Table 7 above (see Appendix A for the full list of stimuli). 80 words that should be known to most L2 speakers and 80 nonwords were selected to serve as targets. For each target, related and unrelated primes were then chosen. Related primes differed only in one letter from the target. Three- or four-letter unrelated primes did not share any letter with the target, and five-letter unrelated primes shared up to one letter with the target. The primes for nonword targets were also words. Finally, it was ensured that all stimuli (both primes and targets) had similar linguistic characteristics (length, log word frequency, neighborhood size and mean bigram frequency) across all prime conditions (see Appendix E for the results of comparisons).

A survey that was administered right after the LDT asked whether participants were able to see the masked primes presented during the LDT. As shown in Table 10, few participants (three L1 speakers and one L2 speaker) felt that they had been able to see more than a half of the primes.

We evaluated the internal consistency of this task. A split-half reliability estimate (Spearman-Brown corrected) suggested that participants' response latency was reliable, r_{SB} = .96 (uncorrected estimates was .93). At the same time, the Rasch person reliability coefficient indicated that accuracy data were acceptable, *r* = .76.

Specific descriptions of the stimuli used for each prime condition follow (see Appendix A for all LDT stimuli).

Table 10

Participants' Perception of Masked Primes

	L1 speakers	L2 speakers
I don't remember something flashed before each target.	26.7%	51.1%
I never identified the primes.	20.0%	26.7%
I rarely identified the primes.	40.0%	16.7%
I identified about a half of the primes.	3.3%	4.4%
I identified more than a half of the primes.	6.7%	1.1%
I always identified the primes.	3.3%	-

Minimal Pairs with the /l/-/r/ Contrast. 16 minimal pairs that consisted of three to five letters (*M* = 3.8 letters) were selected (e.g., *read-LEAD*). In order to trigger greater phonological priming, the member of the pair with the higher frequency (Mean log frequency = 3.67) was always used as the prime, and the other member (Mean log frequency = 2.99) was used as the target (Grainger & Ferrand, 1994). The target had on average 16.6 orthographic neighbors, and the prime had 16.3 neighbors. Unrelated control word primes that did not have a visual form similar to the target were also selected, and the log frequency (*M* = 3.70) and the neighborhood density of the unrelated primes (*M* = 15.1) were matched to those of the related primes.

Minimal Pairs with the /f/-/p/ Contrast. 16 minimal pairs that consisted of three to five letters (*M* = 3.9 letters) were selected (e.g., *full-PULL*). In order to produce stronger phonological priming, the member of the pair with the higher frequency (Mean log frequency = 3.71) was always used as the prime, and the other member (Mean log frequency = 2.88) was used as the target. The target had on average 14.6 orthographic neighbors, and the prime had 13.8 neighbors. Unrelated control word primes that did not have a visual form similar to the target were also selected, and the log frequency (*M* = 3.71) and the neighborhood density of the unrelated primes (*M* = 13.2) were matched to those of the related primes.

Minimal Pairs with the /b/- or /d/-/t/ Contrasts. 16 minimal pairs that consisted of three to five letters (*M* = 3.8 letters) were selected (e.g., *dear-TEAR*). In order to produce stronger inhibition from orthographic neighbors, the member of the pair with the higher frequency (Mean log frequency = 3.87) was always used as the prime, and the other member (Mean log frequency = 3.03) was used as the target (Nakayama et al., 2008; Segui & Grainger, 1990). The target had on average 15.0 orthographic neighbors, and the prime had 14.9 neighbors. Unrelated control word primes that did not have a visual form similar to the target were also selected, and the log frequency (*M* = 3.75) and the neighborhood density of the unrelated primes (*M* = 14.9) were matched to those of the related primes.

Homophone pairs. 16 homophone pairs that consisted of three to five letters (*M* = 3.9 letters) were selected (e.g., *meet-MEAT*). In order to trigger greater phonological priming, the member of the pair with the higher frequency (Mean log frequency = 3.55) was always used as the prime, and the other member (Mean log frequency = 2.81) was used as the target. The target had on average 10.7 orthographic neighbors, and the prime has 13.9 neighbors. Unrelated control word primes that did not have a form similar to the target were also selected, and the log frequency (*M* = 3.59) and the neighborhood density of the unrelated primes (*M* = 13.9) were matched to those of the related primes. Out of 16 homophone pairs, one pair differed only in the first letter (*sell-CELL*), eight pairs differed in one of the median letters (e.g., meet-MEAT), four pairs differed in the last letter (e.g., *pole-POLL*), one pair differed in length (*site-SIGHT*) and two pairs differed in two letters (*one-WON* and *here-HEAR*).

Orthographic Neighbor Pairs that Did Not Differ in the First Letter. 16 pairs that consisted of three to five letters (*M* = 3.9 letters) were selected (e.g., *play-PLAN*). In order to produce stronger inhibition, the member of the pair with the higher frequency (Mean log frequency = 3.95) was always used as the prime, and the other member (Mean log frequency = 3.14) was used as the target. The target had on average 14.8 orthographic neighbors, and the prime had 14.0 neighbors. Unrelated control word primes that did not share any letter with the target were also selected, and the log frequency (*M* = 3.96) and the neighborhood density of the unrelated primes (*M* = 13.5) were matched to those of the related primes. Out of 16 pairs, eight pairs differed in a middle letter (e.g., *type-TAPE*) and eight pairs differed only in the last letter (e.g., *sing-SINK*).

Nonword targets. The selected 80 nonwords consisted of three to five letters (*M* = 3.8) and had on average 6.6 orthographic neighbors when examined by N-Watch (Davis, 2005). These nonword targets (e.g., *shem*) were created by changing one letter from real words that had been selected to serve as their orthographic neighbor primes (e.g., *them*) (Mean log frequency =

3.78, mean neighborhood density = 11.2).[12] Unrelated control word primes that did not have a visual form similar to the target were also selected, and the log frequency (*M* = 3.78), and the neighborhood density of the unrelated primes (*M* = 11.6) were matched to those of the related primes.

For word targets and nonword targets, the stimuli were divided into two groups of prime-target pairs with the two groups having similar lexical characteristics (i.e., word length, frequency, neighborhood density and mean bigram frequency). Two-counterbalanced lists were created so that within each list, each target was paired with one of the two prime types (related vs. unrelated), but across lists, each target was paired with each of the two types of primes.

3.3.2. Spelling Dictation and Spelling Recognition Tasks

The spelling dictation task, which had originally been used by Andrews and her colleagues (Andrews & Hersch, 2010; Andrews & Lo, 2012), asked participants to spell 20 English words they heard. More specifically, participants first heard a target word and then a sentence that helped them know the context where the word was used. Participants then typed in the spelling of the target they heard. According to Rasch person reliability, the internal consistency of the task (k = 20) for the participant sample (N = 120) was .81.

[12] Although all nonwords were created to be pronounceable, orthotactic regularities in English were not seriously considered when creating nonword foils (e.g., *UCT*) in the current study. This might have affected the results from participants who had sufficient experience of reading in English and possessed implicit knowledge of English orthotactic rules. A suspected impact of this type of nonword is relatively easier discrimination of word targets from nonword foils that might lead to weaker lexical competition. However, orthotactic regularities have rarely been intensely controlled in most of the previous studies reviewed in Section 2.1, so we assume that the results obtained in the present study were produced with the materials similar to those in previous studies on orthographic neighbor priming.

The spelling recognition task consisted of 88 items, half spelled correctly and half misspelled. Using Qualtrics, the list of items was displayed on computer screen, with binary choice buttons "Correct" and "Wrong" presented next to each item. Participants were then asked to choose one of the two options (See Appendix B for all items that were used in the spelling dictation and spelling recognition tasks). According to Rasch person reliability, the internal consistency of the task (k = 88) for the participant sample (N = 120) was .86.

3.3.3. Vocabulary test

To measure participants' vocabulary size, a vocabulary test which was developed by Shipley (1940) was employed. Brysbaert, Lagrou and Stevens (2017) used this vocabulary test for their study on visual word recognition and noted that this test was a *good* measure of vocabulary size, a fine proxy for language exposure, and they recommend all language researchers to use such a test (p. 545). This test consisted of 40 items, and participants were asked to choose a word that had a similar meaning to a given word (e.g., *TALK*) out of four options (e.g., a. *draw*, b. *eat*, c. *speak*, d. *sleep*) (see Appendix C for all items). According to Rasch person reliability, the internal consistency of the test (k = 40) for the participant sample (N = 120) was .74.

3.3.4. Word and Phoneme Identification Tasks

Ninety-six stimuli were used for the word identification task. Each of the 96 words that were used either as primes or targets under the three critical prime conditions of the LDT (i.e., the conditions for the /l/-/ɹ/, /p/-/f/, and /b/- or /d/-/t/ contrasts) were presented. Ninety-six stimuli were also used for the phoneme identification task. These stimuli were made by either changing the last phoneme of a word stimulus used for the word identification task (e.g., *read*→/ɹiθ/) or adding another phoneme at the end of a word stimulus (e.g., *ray*→/ɹeɪv/) (See Appendix B for the items). Through applying this rule for creating pseudoword stimuli, the only difference between the stimuli used for the word identification task and those used for the phoneme identification task was the fact that the former were words whereas the latter were pseudowords, as all critical phonemes (e.g., /ɹ/) in both tasks were followed by the same vowels (e.g., /i/ or /eɪ/) across the two tasks. The above-described stimulus selection procedure for the word and phoneme identification tasks made these two tasks qualitatively different from the measures of orthographic precision (spelling tasks), vocabulary size (a vocabulary test) and global proficiency (a cloze test). Specifically speaking, whereas the other measures attempted to assess the levels of each target construct using a set of materials that did not target any specific linguistic features or difficulties, the word and phoneme identification tasks directly assessed participants' ability to accurately perceive the critical contrasts that had appeared in the LDT.

For each of these two tasks, two presentation lists were prepared. In the first half of the first list, one member of each phonological contrast (e.g., *read*) was presented first, and the other member of the contrast (e.g., *lead*) was presented in the second half. The presentation order of the two members of the contrast was reversed in the second list (e.g., *lead* was presented in the first half and *read* was presented in the second half). Half of the participants were provided with the first list, and the other half were presented with the second list.

In both sound perception tasks, participants sat in front of a computer. They heard each stimulus pronounced at a natural speed by a male native English speaker only one time. In the word identification task, after hearing the stimulus, they were asked to choose a word (e.g., *read*) that they had just heard out of two options (a minimal pair such as *read* vs. *lead*). In the phoneme identification task, participants heard each pseudoword stimulus and were asked to choose the first phoneme of the stimulus that they had just heard (e.g., /riθ/) out of two options (a phonological contrast visually presented in English letters such as /l/ and /r/). In both tasks, a plus mark (+) was first presented at the center of the screen, and then an aural stimulus followed. After that, two options were visually presented on the screen, and participants' response time and accuracy on each stimulus were recorded. The onset of the reaction time was the moment when two options appeared on the screen after the presentation of an aural stimulus. Participants first responded to 10 practice stimuli with feedback on each of their responses, but no feedback was provided for critical stimuli (See Appendix D for all items that were used in the word identification and phoneme recognition tasks).

The internal consistency of the two sound perception tasks was evaluated. The Rasch person reliability coefficient for accuracy data from the word identification task was 61. The Rasch person reliability coefficient for accuracy data from the phoneme identification task was 70. These relatively low Rasch person reliability coefficients were produced because both L1 and L2 participants performed very well on control stimuli (words or nonwords with the /b/-/t/ or /d/-/t/ contrast). In other words, almost every participant responded to all of those stimuli correctly, so the task was not very meaningful for the purpose of distinguishing the participants who were better and worse at accurate phonological categorization of the /b/-/t/ and /d/-/t/ contrasts.[13]

[13] The Cronbach Alpha coefficient for the word identification task was .88, and that for the phoneme identification task was .87. These high coefficient values suggest that in the two tasks, participants did not respond to items randomly, and therefore, the measures were reliable. However, we were more interested whether these two tasks were reliably measuring a target *construct* (phonological categorization) rather than whether the scores from a particular task were repeatable. Therefore, because Rasch person reliability provides more valid information about whether *a test produces repeatable measures for a sample* (recruited participants) while taking the difficulty of the test into consideration, whereas a Cronbach Alpha coefficient only indicates the repeatability of raw scores from a single test regardless of whether the test is good or not for measuring a target construct (Linacre, 1997), Rasch reliability was used in the present study.

However, in data analysis, we only used participants' responses to critical stimuli (words and nonwords with the /f/-/p/ and /l/-/ɹ/ contrasts) which showed an acceptable level of Rasch person reliability (see below). Put differently, the results we report in the next chapter were based on data that were sufficiently reproducible.

3.3.5. Cloze Test

A cloze test (Brown, 1980) was administered to measure participants' proficiency. In this task, participants were asked to fill in 50 blanks distributed throughout a reading passage while reading within a limited time (20 minutes). Only L2 participants took this test (see Appendix E for the test). According to Rasch person reliability, the internal consistency of the test (k = 50) for the participant sample (N = 88) was .88.

3.3.6. Translation and Familiarity Rating Tasks

L2 speakers were asked to provide both a familiarity rating and an L1 translation equivalent for each LDT stimulus. They were also requested to translate all the stimuli used in the spelling dictation and recognition tasks. On the other hand, L1 speakers were asked to give familiarity ratings for all the spelling dictation and recognition stimuli. What follow are the reasons for implementing these tasks and the descriptions of them.

For L2 speakers, unknown words are supposed to function like nonwords. At the same time, nonword primes are known to produce facilitation (e.g., Forster & Veres, 1998, Lupker & Davis, 2006), and any priming effects for nonword targets are not clearly interpretable. Therefore, to minimize the influence of unknown words, a translation and familiarity rating task was administered to ask whether there had been any word stimuli in the LDT that was unknown to the L2 participants.

In this task, L2 speakers were requested to give a familiarity rating for each LDT stimulus along with its translation. For familiarity ratings, three options were provided: "I have never seen this word before," "I know that this is an English word" and "I know the meaning of this word." In Park's study (2021), only word stimuli that were reported to be familiar were included in data analysis because the purpose of the study was to investigate whether there would still be no lexical competition even when both the prime and the target were familiar. However, the purpose of the present study was to observe any changes in the patterns of form priming across the three participant groups, and it was reasonable to assume that the L2 lexicon could sometimes include unfamiliar words (i.e., words whose forms were known but their meanings were unknown). Therefore, in the present study, only stimuli with "I have never seen this word before" responses were excluded from data analysis to exclude the effects of totally unknown words but not of less-familiar words. At the same time, participants' translations were used for qualitative analysis of participants' orthographic precision.

L2 speakers were also asked to translate the words that had appeared in the spelling dictation and spelling recognition tasks. The previous studies on the role of orthographic precision in lexical competition between L1 words or in L1 reading comprehension (Andrews et al., 2020; Andrews & Hersch, 2010; Andrews & Lo, 2012) treated spelling knowledge (i.e., the knowledge of the spellings of familiar words) as a construct that was distinguished from vocabulary knowledge (i.e., the knowledge of the meanings of rarely used words). Since this task was to measure spelling knowledge but not vocabulary knowledge or proficiency, the L2 words for which incorrect L1 translation equivalents were provided were excluded from analysis. At the same time, L1 speakers also indicated their familiarity with each item that was presented in the two spelling tasks. In this familiarity task, three familiarity ratings were given: “I have never seen this word before,” “I know that this is an English word” and “I know the meaning of this word.” The words for which the “I have never seen this word before” or “I know that this is an English word” responses were selected were then excluded from analysis. By following these procedures, the spelling tasks assessed the spellings of words whose meanings were known to participants.

Participants’ L1 translations of English word stimuli used in the LDT were separately scored by the researcher (a native speaker of Korean) and another native Korean L2 learner of English, but translations of the items used in the spelling dictation and the spelling recognition tasks were scored only by the researcher. When scoring, a response was considered correct if it showed a participant’s correct understanding of the lexeme of each item regardless of its inflections or his/her spelling error in its L1 translation. Percentage agreement between the two raters was 98.08% (Cohen’s Kappa was .90). Where disagreements occurred, the researcher made the final decision.

3.4. Procedure

All experiments were conducted online. The overall procedure of the current study is presented in Figure 2.

When participants met the researcher on Zoom, they first performed the masked priming LDT that was programmed and displayed in PsychoPy. During this task, participants were asked to decide whether a stimulus they saw on the screen was an English word, and to press the “j” (for a “yes” response”) or the “f” key (for a “no” response) as accurately and quickly as possible, but not so quickly that they made too many errors (Forster & Veres, 1998). At the beginning of each trial, the mask "#####" was first presented (for 500 ms) at the center of the screen. The prime (either a related or an unrelated prime) was then displayed for 67 ms and followed by the target (either a word or nonword). This prime duration, which was a little longer than the ones used in typical masked priming experiments (40-60 ms) was expected to promote stronger inhibition from orthographic neighbor primes and stronger facilitation from phonological primes (Ferrand & Grainger, 1992), but was not long enough for the prime to become visible (Nakayama & Lupker, 2018) (see Table 10 for the visibility of the primes as perceived by participants). The LDT was preceded by 10 practice trials and 6 "buffer" trials, among which there were equal numbers of word and nonword stimuli. These trials were not used for data analysis. Items that were used for data analysis (i.e., 80 word and 80 nonword targets) were presented in a random sequence, with a different order for each participant.

After the LDT, both L1 and L2 participants carried out the spelling (dictation and recognition) tasks, vocabulary test and sound perception tasks (word and phoneme identification tasks). Participants carried out the spelling tasks first before performing the word identification task to minimize the influence of their prior encounters with word targets during the LDT (Note that all items that were used in the word identification task were from the stimuli used in the LDT). L2 participants also took the cloze test and performed translation and familiarity tasks that asked their familiarity with the words that had appeared in the LDT and two spelling tasks. On the other hand, L1 participants carried out a familiarity task that asked their familiarity with spelling-task stimuli. It took about 50 minutes for L1 participants (range: 40 ~ 60 minutes) and 100 minutes for L2 participants (range: 90 ~ 150 minutes) to complete the whole experiment.

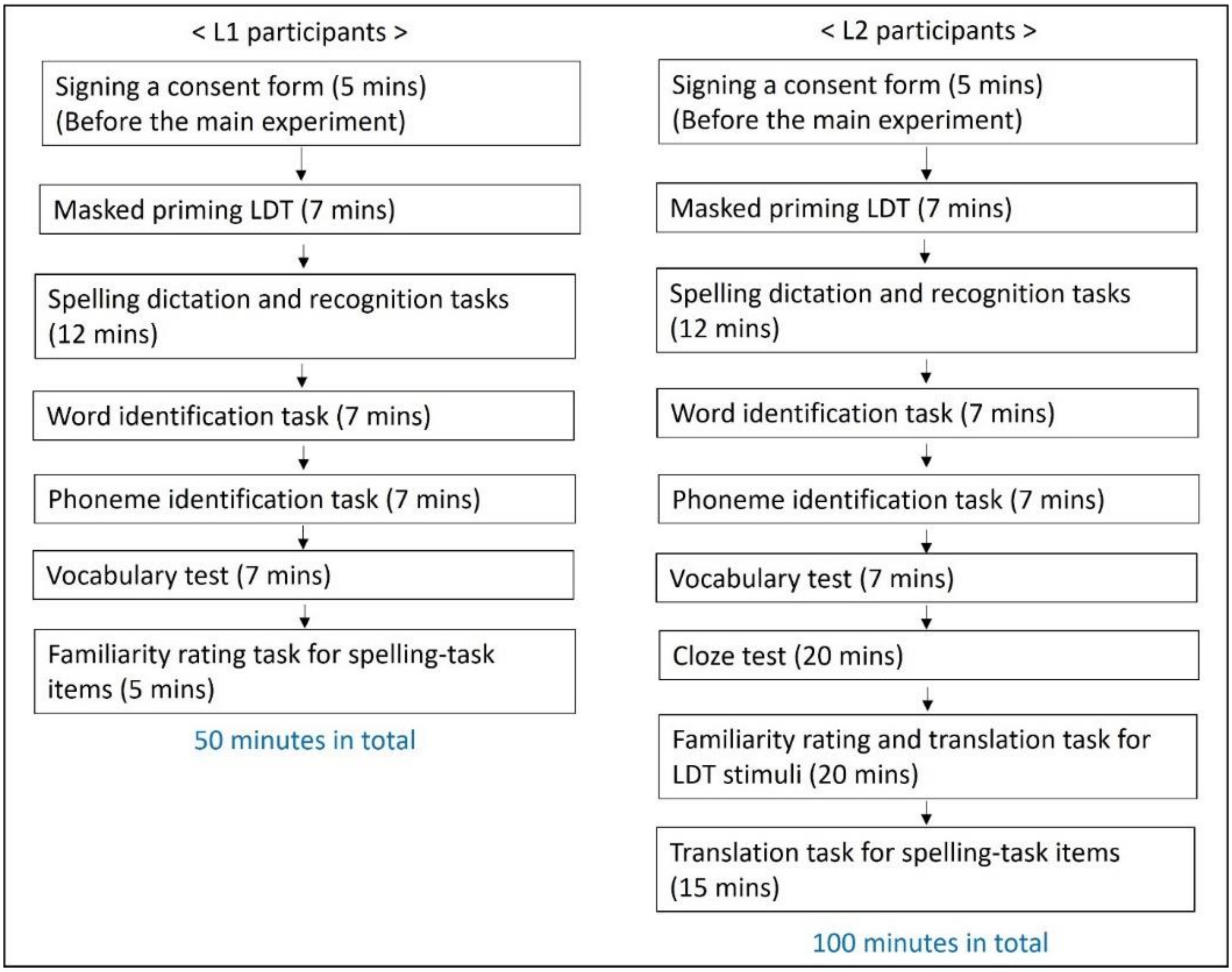

Figure 2. The overall procedure and average time spent for each task.

Chapter 4. Results

4.1. Sound Identification of Phonological Contrasts

The first research question of the present study asked whether Korean L2 learners of English have trouble identifying the sounds of two nonnative phonological contrasts, the /l/-/ɹ/ and /f/-/p/ contrasts. This research question was important for the current study because it investigates the role in visual word processing of fuzzy phonological representations of words with these two contrasts. Thus, to estimate the quality of phonological representations, we used a word identification task. We reasoned that if L2 speakers had established accurate phonological representations in their mental lexicon, this would be evident in their performance in the word identification task. This was not only because if they could accurately perceive aurally presented words in L2 input, their perception would help them develop precise phonological representations (e.g., Pallier et al., 2001), but also because precise phonological representations might help their accurate perception in reverse (e.g., Flege, 2003; Meador et al., 2003; Thorn & Gathercole, 2001). In addition to the word identification task, a phoneme identification task was used to examine whether the accuracy of L2 sound identification would differ depending on the lexical status of the stimuli.

For most statistical analyses reported in this chapter, data from three L2 participant groups (low-, medium- and high-proficiency L2 speakers) were combined. This provided us with a group of L2 speakers with a wide range of proficiency.

When participants carried out the word and phoneme identification tasks, their reaction times were recorded. These reaction time data (not reported in this paper) were used for data trimming. When responses were 3 standard deviations (SD) faster or slower than each individual participant’s mean reaction time, these responses were excluded from analyses since they could have been affected by something unexpected, such as loss of participants' attention during the experiment. This resulted in exclusion of 2.1% of data.

The accuracy rates of L1 and L2 speakers' sound identification performance are presented in Table 11 and Figure 3. Most L2 speakers, as well as L1 speakers, obtained a perfect score for identification of contrasts that were differentiated in their L1 (the /b/-/t/ and /d/-/t/ contrasts). As noted earlier, the Rasch person reliability coefficient for the word identification task was low (.61) because the task was easy. This was particularly the case for the control stimuli (i.e., the stimuli with the /b/-/t/ and /d/-/t/ contrasts). The internal consistency of participants' responses to the control stimuli (which were combined from both the word and the phoneme identification tasks) was extremely low (.00), indicating that the identification of the control stimuli was too easy, so the same results would not be reproducible. Therefore, this paper does not report the results of statistical tests for control items.

Table 11

L1 and L2 Speakers' Accuracy Rates in Word and Phoneme Identification

		/b/ - /t/	/d/ - /t/	/f/ - /p/	/l/ - /ɹ/
Word	L1 speakers	.992 (*SD* = .092)	.983 (*SD* = .129)	.985 (*SD* = .121)	.984 (*SD* = .125)
	L2 speakers	.998 (*SD* = .046)	.991 (*SD* = .095)	.964 (*SD* = .185)	.873 (*SD* = .333)
Nonword	L1 speakers	1.00 (*SD* = .00)	.989 (*SD* = .102)	.989 (*SD* = .103)	.971 (*SD* = .167)
	L2 speakers	.999 (*SD* = .065)	.999 (*SD* = .046)	.959 (*SD* = .199)	.822 (*SD* = .385)

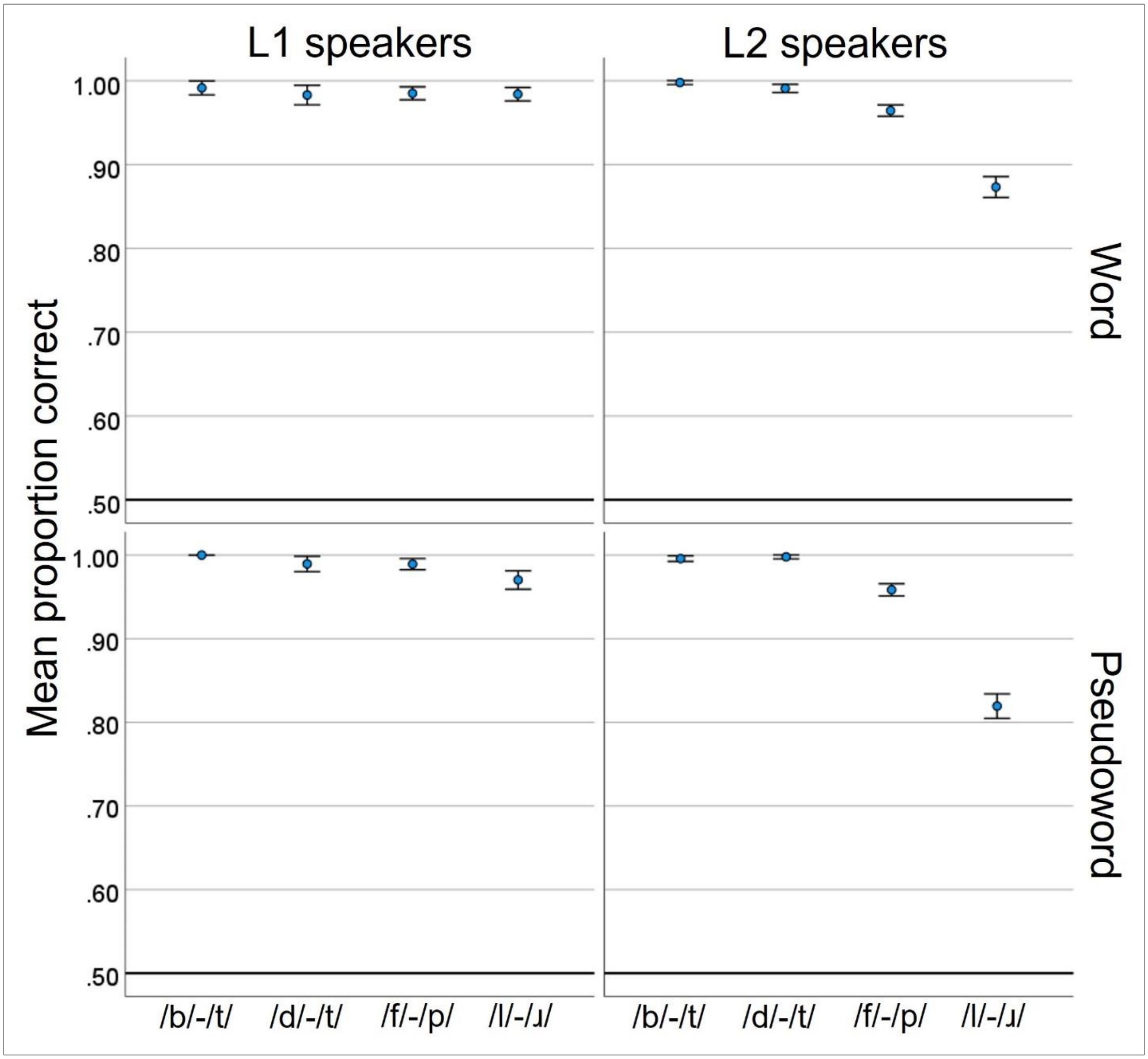

Figure 3. Accuracy Rates for Word Stimuli (Top) and Nonword Stimuli (Bottom) as a Function of Language Group (L1 (Left) vs. L2 (Right)), and Contrast (/b/–/t/, /d/-/t/, /f/-/p/ and /l/-/ɹ/).

Rasch person reliability indicated that the internal consistency of participants' responses to critical stimuli (that were combined from the word and the phoneme identification tasks) was acceptable (.70). Thus, further analysis was performed for the critical stimuli using a mixed-effects model. In this model, response accuracy (coded as ‘0 = incorrect’ and ‘1 = correct’ for each trial) was the dependent variable, and Group (L1 vs. L2 speakers), Contrast (the /f/-/p/ vs. /l/-/ɹ/ contrasts), Lexical Status (word vs. nonwords) and their interactions were fitted as fixed effects. All the fixed effects were coded as ‘0’ or ‘1’. Participant and Item were fitted as random effects. We also examined whether adding a random slope for each fixed effect or for an interaction between them improved model fit. Because adding a random slope for Group to the random effect Participant significantly improved model fit, this random slope was included. The outputs of the model are shown in Table 12. These outputs are all from a single model, but to fully understand all the effects of the predictors under each of eight reference categories, the model was examined again for each reference category.

Identification of words with the /f/-/p/ contrast was not significantly worse for L2 speakers than L1 speakers ($\beta = -0.30$, $SE = 0.33$, $z = -0.92$, $p = .36$). However, L2 speakers were significantly less accurate in identification of the /f/-/p/ contrast when those sounds were embedded in nonwords ($\beta = -0.82$, $SE = 0.37$, $z = -2.23$, $p = .03$). This indicates that unlike the /b/-/t/ and /d/-/t/ contrasts for which most L2 speakers obtained a perfect score, identification of the nonnative /f/-/p/ contrast was indeed challenging to L2 speakers, at least in nonwords.

Table 12

L1 and L2 Speakers' Accuracy Rates in Word and Phoneme Identification

Fixed effects				
Predictor	Estimate	*SE*	*z*	*p*
(1) Intercept: L1 speakers, /f/-/p/, Word	4.42	0.31	14.49	< .001***
L2 speakers	-0.30	0.33	-0.92	.356
/l/-/ɹ/	-0.02	0.41	-0.06	.953
Pseudoword	0.37	0.44	0.84	.400
L2 speakers * /l/-/ɹ/	-1.52	0.39	-3.93	< .001***
L2 speakers * Pseudoword	-0.52	0.43	-1.20	.231
/l/-/ɹ/ * Pseudoword	-0.91	0.57	-1.59	.111
L2 speakers * /l/-/ɹ/ * Pseudoword	0.55	0.55	1.02	.308
(2) Intercept: L1 speakers, /f/-/p/, Pseudoword	4.79	0.35	13.79	< .001***
L2 speakers	-0.82	0.37	-2.23	.026*
/l/-/ɹ/	-0.94	0.40	-2.31	.021*
L2 speakers * /l/-/ɹ/	-0.97	0.39	-2.51	.012*
(3) Intercept: L1 speakers, /l/-/ɹ/, Word	4.40	0.30	14.60	< .001***
L2 speakers	-1.83	0.32	-5.76	< .001***
Pseudoword	-0.54	0.37	-1.48	.139
L2 speakers * Pseudoword	0.04	0.33	0.13	.900
(4) Intercept: L1 speakers, /l/-/r/, Pseudoword	3.86	0.25	15.34	< .001***
L2 speakers	-1.79	0.27	-6.69	<. 001***
(5) Intercept: L2 speakers, /f/-/p/, Word	4.12	0.21	19.53	< .001***
/l/-/ɹ/	-1.54	0.20	-7.59	< .001***
Pseudoword	-0.14	0.22	-0.66	.506
/l/-/ɹ/ * Pseudoword	-0.35	0.28	-1.25	.307
(6) Intercept: L2 speakers, /f/-/p/, Pseudoword	3.97	0.21	19.09	< .001***
/l/-/ɹ/	-1.90	0.20	-9.55	< .001***
(7) Intercept: L2 speakers, /l/-/ɹ/, Word	2.57	0.19	13.45	< .001***
Pseudoword	-0.50	0.18	2.73	.006**
Random effects	Variance	SD	Corr	
Participant	0.15	0.40		
Participant-Group	1.33	1.15	0.14	
Item	0.41	0.64		

Notes. When the effect of a predictor was identified in an earlier examination and was once noted in the table, the same effect that was identified after switching the reference category was not noted again since it was redundant. For example, when the effect of a three-way interaction between three predictors was typed in the third row of the table, the same information that was identified after changing the reference category was not retyped in the fourth row. Formula (glmer): Acc ~ Group * Condition * Lexicality + (1 + Group | Participant) + (1 | Item), Significance codes: '***' < .001, '**' < .01, '*' < .05

L2 speakers performed significantly worse than L1 speakers in identification of the /l/-/ɹ/ contrast, no matter the stimuli were words ($\beta = -1.83$, $SE = 0.32$, $z = -5.76$, $p < .001$) or nonwords ($\beta = -1.79$, $SE = 0.27$, $z = -6.69$, $p < .001$). This finding is consistent with previous research that reported native Korean speakers' difficulty in acquisition of the /l/-/ɹ/ contrast (Borden et al., 1983; Jamieson & Yu, 1996; Smith, 2001). The /l/-/ɹ/ contrast was also less accurately identified by L2 speakers than the /f/-/p/ contrast both in words ($\beta = -1.54$, $SE = 0.20$, $z = -7.59$, $p < .001$) and nonwords ($\beta = -1.90$, $SE = 0.20$, $z = -9.55$, $p < .001$). This finding supports the prediction that a nonnative contrast in which two L2 phonemes are perceived to be equally distant from a relevant L1 phoneme is more difficult to learn than another contrast in which one L2 phoneme is perceived to be closer to a relevant L1 phoneme than the other phoneme (Best & Tyler, 2007; Flege, 2003). As noted earlier, Schmidt (1996) reported that the native Korean participants in her study perceived that /l/ and /ɹ/ were almost equally distant from a Korean consonant 'ㄹ', whereas they felt that /p/ was closer to a Korean consonant 'ㅍ' than /f/.

Finally, the effect of the lexicality of stimuli was not significant either for the /f/-/p/ or /l/-/ɹ/ contrasts in L1 speakers. It was also non-significant for the /f/-/p/ contrast in L2 speakers, but lexicality was significant for the /l/-/ɹ/ contrast ($\beta = -0.50$, $SE = 0.18$, $z = 2.73$, $p = .006$). This finding is in line with the hypothesis that the acquisition of the phonological forms of words may be able to precede the acquisition of phonological categories (e.g., Darcy et al., 2012; Gor, 2015, 2018).

To sum up, the auditory tasks confirmed that L2 participants found it more difficult to identify the /l/-/ɹ/ contrast than L1 participants. Moreover, even though identification of the /f/-/p/ contrast was found to be easier for L2 speakers than that of the /l/-/ɹ/, this contrast was still significantly less accurately identified by L2 speakers compared to L1 speakers if it was embedded in nonwords.

4.2. Orthographic Neighbor Priming in L1 and L2 Speakers

Subsequent analysis was performed to compare the patterns of form priming for L1 and L2 speakers. Responses to words and nonwords were examined separately when analyzing data from the LDT. Before the experiment, the plan was to remove the data from participants whose error rate was higher than 25%. However, no participant exceeded that threshold level, so no data were excluded. We dealt with outliers by establishing cut-offs at 2.5 SD above or below each participant's mean reaction time. One response to word targets and two responses to nonword targets with a response latency shorter than 300 ms were also excluded. This resulted in the loss of 2.0% of L1 speakers' and 0.9% of L2 speakers' responses to word targets, and 3.6% of L1 speakers' and 3.3% of L2 speakers' responses to nonword targets. To filter out any influence of unknown words from L2 speakers' responses to word targets, 1.7% of responses that involved either unknown primes or targets were also excluded.[14] These were 26 responses to targets primed by unknown unrelated primes, 9 responses to targets primed by unknown related primes, 43 unknown targets primed by unrelated primes, and 56 unknown targets primed by related primes. When analyzing response latency, 5.6% of L1 speakers' and 8.0% of L2 speakers' inaccurate responses to word targets, and 4.5% of L1 speakers' and 6.7% of L2 speakers' inaccurate responses to nonword targets were further excluded.

[14] As noted above, unknown stimuli were identified by a translation and familiarity rating task.

For statistical analyses, linear mixed-effect models (for response latency analyses) and logistic mixed-effects models (for accuracy analyses) were fitted using lme4 package (Bates, Maechler, Bolker, & Walker, 2013) in R. In models for logistic regression analyses, the dependent variable was correctness of the response to the target (coded as '0 = incorrect' and '1 = correct' for each trial), and in models for linear regression analyses, the dependent variable was inverse-transformed correct response latencies to the target. Participant and Item were fitted as random effects. Prime type (coded as '-0.5 = unrelated' and '+0.5 = related'), Prime Condition (coded as '1', '2', '3', '4' and '5' for each condition) and Group (coded as '0 = L1 speakers' and '1 = L2 speakers') were fitted as fixed effects. The same model building procedure that was described in Section 4.1 was followed.

Mean reaction times and error rates for word and nonword targets under different prime conditions are presented in Table 13.

4.2.1. Response Latency for Word Targets

The outputs of a mixed effect model are presented in Table 14. The same model was examined switching the reference category for Group. When the reference category was L1 speakers, the effects of Prime Type (related vs. unrelated primes) (β = 39.03, SE = 27.92, t = 1.40, p < .162) and Prime Condition (five prime conditions) (β = -2.82, SE = 7.58, t = -0.37, p = .710) were nonsignificant. However, the effect of Group was significant (β = 239.77, SE = 46.90, t = 5.11, p < .001) indicating that L2 speakers responded to targets significantly more slowly than L1 speakers. The interaction between Prime Type and Group was also significant (β = −114.71, SE = 29.38, t = -3.91, p < .001), suggesting that the effect of Prime Type differed between L1 and L2 speakers. In addition, there was a significant three-way interaction between

Prime Type, Prime Condition and Group (β = 22.85, *SE* = 8.82, *t* = 2.59, *p* = .010), indicating that the interaction between Prime Type and Prime condition significantly differed between the two participant groups. On the other hand, the interactions between Prime Type and Prime Condition and between Prime Condition and Group were non-significant.

When the reference category for Group was changed to L2 speakers, the effect of Prime Type was significant (β = -75.68, *SE* = 19.34, *t* = -3.91, *p* < .001), suggesting that form facilitation was significant in L2 speakers. There was also a significant interaction between Prime Type and Prime Condition (β = 17.92, *SE* = 5.79, *t* = 3.09, *p* = .002). This interaction means that the patterns and the strength of form priming were different across the five prime conditions.

To understand the interaction between Prime Type and Prime Condition, Prime Condition was refitted as a factor variable in the same model used earlier. Then, while changing the reference category for Prime Condition, the main effect of Prime Type under each Prime Condition was examined. Although the interaction between Prime Type and Prime Condition was non-significant in L1 speakers, the same examination was carried out for them for comparison with L2 speakers. Table 15 presents the results of this examination.

Table 15 shows that form facilitation for homophone pairs (β = -47.18, *SE* = 18.75, *t* = -2.52, *p* = .012) and minimal pairs with the contrast /l/-/r/ was significant (β = -50.79, *SE* = 18.82, *t* = -2.70, *p* = .007) in L2 speakers. On the other hand, form priming under the other prime conditions was not significant in this group of participants. At the same time, form inhibition was non-significant under all prime conditions in L1 speakers.

Table 13

L1 and L2 Speakers' Mean Reaction Times (in Milliseconds) and Error Rates (in Parentheses) under Different Prime Conditions in the Masked Priming Lexical Decision Task

	Homophone		/l/ - /ɹ/		/p/ - /f/		/d/ - /t/ or /b/ - /t/		Orthographic neighbors sharing the word initial		Nonwords	
	Related (*meet-MEAT*)	Unrelated (*kids-MEAT*)	Related (*read-LEAD*)	Unrelated (*wish-LEAD*)	Related (*full-PULL*)	Unrelated (*sick-PULL*)	Related (*dear-TEAR)*	Unrelated (*boys-TEAR)*	Related (*play-PLAN*)	Unrelated (*best-PLAN*)	Related (*name-MAME*)	Unrelated (*long-MAME*)
L1 speakers (*n* = 30)	657.0 (5.1%)	640.2 (4.7%)	685.4 (9.7%)	665.8 (5.6%)	671.7 (6.8%)	655.2 (5.2%)	657.0 (4.7%)	649.2 (3.8%)	655.8 (6.4%)	655.8 (4.2%)	700.8 (4.6%)	704.7 (4.4%)
	-16.8 (-0.4%)		-19.6† (-4.1%)		-16.5 (-1.6%)		-7.8 (-0.9%)		0 (-2.2%)		3.9 (-0.2%)	
L2 speakers (*n* = 90)	780.5 (8.6%)	801.2 (12.3%)	788.9 (11.4%)	809.5 (11.8%)	791.8 (6.8%)	805.6 (7.1%)	762.9 (6.6%)	771.6 (8.7%)	807.5 (6.3%)	790.2 (7.2%)	856.2 (7.2%)	862.6 (6.2%)
	20.7* (3.7%)		20.6** (0.4%)		13.8 (0.3%)		8.7 (2.1%)		-17.3 (0.9%)		6.4 (-1.0%)	

Notes. Significance codes: '**' < .01, '*' < .05

Table 14

Output from a Linear-Mixed Effects Model for Inverse-Transformed Reaction Time Data (RTs) in L1 and L2 Speakers' Masked Priming Lexical Decision

Fixed effects				
Predictor	Estimate	*SE*	*t*	*p*
(1) The reference category for Group: L1 speakers				
(Intercept)	-1595.78	46.05	-34.65	< .001***
Related	39.03	27.92	1.40	.162
Condition	-2.82	7.58	-0.37	.710
GroupL2	239.77	46.90	5.11	< .001***
Related:Condition	-4.93	8.41	-0.59	.558
Related:GroupL2	-114.71	29.38	-3.91	< .001***
Condition:GroupL2	2.94	4.42	0.67	.505
Related:Condition:GroupL2	22.85	8.82	2.59	.010**
(2) The reference category for Group: L2 speakers				
(Intercept)	-1356.01	32.07	-42.29	< .001***
Related	-75.68	19.34	-3.91	< .001***
Condition	0.12	6.95	0.02	.986
Related:Condition	17.92	5.79	3.09	.002**
Random effects	Variance	SD	Corr	
Participant	44629	211.26		
Item	6898	83.06		
Related	2104	45.87	-0.15	

Notes. Formula (lmer): RT_Inverse ~ Related * Condition * Group + (1 | Participant) + (1 + Related | Item), Significance codes: '***' < .001, '**' < .01, '*' < .05

In short, the current study has replicated the findings of previous research in which L1 and L2 speakers showed different patterns of orthographic priming (Nakayama & Lupker, 2018; Park, 2021; Qiao & Forster, 2017, Jiang, 2021)[15]. However, it was additionally observed that

[15] After the experiment, it was found that the LDT stimuli included three heteronyms, words that have a different pronunciation and meaning from another word but the same spelling (*read*, *lead* and *tear*). For that reason, the results were reexamined after excluding two prime target pairs (*read-LEAD* and *dear-TEAR*) because these heteronyms could have affected the results in an unexpected way. Overall, the results of this additional analysis were very similar to those presented in Tables 13 and 16. Interested readers can refer to Appendix I for these results.

when the form priming under each prime condition was examined separately, L2 speakers showed significant facilitation for homophone pairs and minimal pairs with the /l/-/ɹ/ contrast but not for other orthographic neighbor prime-target pairs. This latter finding has not been reported in previous research. The reasons for and the implications of this phenomenon are discussed below.

Table 15

Form Priming under Five Prime Conditions Observed in L1 and L2 Speakers

Group	Prime Condition	Form priming	Estimate	*SE*	*t*	*p*
L1 speakers	Homophone	-16.8 ms	21.88	26.68	0.82	.412
	/l/-/ɹ/	-19.6 ms	47.74	27.06	1.76	.078†
	/f/-/p/	-16.5 ms	18.52	26.91	0.69	.491
	/b/-/t/ or /d/-/t/	-7.8 ms	28.59	26.63	1.07	.283
	Pairs sharing the initial	0 ms	5.39	26.78	0.20	.841
L2 speakers	Homophone	20.7 ms	-47.18	18.75	-2.52	.012*
	/l/-/ɹ/	20.6 ms	-50.79	18.82	-2.70	.007**
	/f/-/p/	13.8 ms	-21.26	18.48	-1.15	.250
	/b/-/t/ or /d/-/t/	8.7 ms	-15.60	18.49	-0.84	.399
	Pairs sharing the initial	-17.3 ms	24.93	18.36	1.36	.175

Notes. Significance codes: '**' < .01, '*' < .05, †: $p < .10$

4.2.2. Accuracy on Word Targets

The outputs of a model used to analyze accuracy for word targets are presented in Appendix H. L1 speakers were significantly more accurate in lexical decisions than L2 speakers

(β = -0.87, *SE* = 0.28, *z* = -3.16, *p* = .002). However, the effects of Prime Type and Prime Condition as well as their interactions were all non-significant.

4.2.3. Response Latency to Nonword Targets

The outputs of a model used to analyze response latency to nonword targets are presented in Appendix H. L2 speakers' responses were significantly slower than L1 speakers' (β = -245.49, *SE* = 46.21, *t* = -5.31, *p* < .001). However, the effects of Prime Type and of the interaction between Prime type and Prime Condition were non-significant.

4.2.4. Accuracy for Nonword Targets

The outputs of a model used to analyze accuracy for nonword targets are presented in Appendix H. The effects of Prime Type, Group and their interactions were all non-significant.

4.3. Patterns of Form Priming for Three Groups of L2 Speakers

Research Question 3 of the present study asked whether different patterns of form priming are observed across three groups of L2 speakers with different levels of proficiency. Therefore, the pattern of priming for each of the three L2 speaker groups was examined separately. The descriptive statistics for this analysis are presented in Table 16. For easier comparison with L1 speakers' data, this group's results are also included.

For statistical analysis, the same model construction procedure that was described in Section 4.2 was followed except that the fixed effect Group was not included in models. A summary of the results of statistical tests (whether each priming is significant or not) for nonword targets is also presented in Table 16. However, these results will not be specifically described or discussed in the main text because they are not closely related to the research questions.

4.3.1. Low-proficiency L2 Speakers

Response Latency[16]. Whereas the main effect of Prime condition was not significant (β = -0.15, *SE* = 8.38, *t* = -0.02, *p* = .986), the main effect of Prime Type was significant (β = -110.48, *SE* = 27.55, *t* = -4.01, *p* < .001) indicating that there was significant facilitation across the prime conditions. There was also a significant interaction between Prime Type and Prime Condition (β = 26.95, *SE* = 8.18, *t* = 3.30, *p* = .001). This suggests that the patterns or the strength of form priming varied depending on the prime condition.

To explore the nature of the interaction between Prime Type and Prime Condition, whether form facilitation was significant under each prime condition was examined while switching the reference category for Prime Condition. The results are presented in Table 17. In short, form facilitation was significant only under the homophone (β = -69.32, *SE* = 26.57, *t* = -2.61, *p* = .009) and the /l/-/ɹ/ contrast conditions (β = -60.14, *SE* = 26.45, *t* = -2.27, *p* = .023) but not under the other prime conditions.

Accuracy[17]. Neither was the main effect of Prime Type (β = 0.15, *SE* = 0.36, *z* = 0.43, *p* = .667) nor that of Prime Condition (β = 0.24, *SE* = 0.13, *z* = 1.76, *p* = .078) significant. The interaction between Prime Type and Prime Condition (β = 0.06, *SE* = 0.12, *z* = 0.48, *p* = .635) was also non-significant.

[16] Model formula (lmer): RT_Inverse ~ Related * Condition + (1 | Participant) + (1 | Item)
[17] Model formula (lmer): Acc ~ Related * Condition + (1 | Participant) + (1 | Item)

Whether form facilitation was significant under each prime condition was also inspected. As shown in Table 16, facilitation was only significant under the /b/-/t/ or /d/-/t/ contrast condition (β = 1.29, *SE* = 0.49, *z* = 2.63, *p* = .008).

4.3.2. Medium-proficiency L2 Speakers

Response Latency[18]. Whereas the main effect of Prime condition was not significant (β = 0.83, *SE* = 9.03, *t* = 0.09, *p* = .927), the main effect of Prime Type was significant (β = -93.61, *SE* = 26.95, *t* = -3.47, $p < .001$) indicating that there was significant facilitation across the prime conditions. There was also a significant interaction between Prime Type and Prime Condition (β = 22.57, *SE* = 8.42, *t* = 2.68, *p* = .007). This suggests that the patterns or the strength of form priming differed depending on the prime condition.

Whether form facilitation was significant under each prime condition was also inspected. The results are presented in Table 16. In short, form facilitation was significant only under the homophone (β = -58.50, *SE* = 25.06, *t* = -2.34, *p* = .020) and the /l/-/ɹ/ contrast conditions (β = -63.37, *SE* = 24.97, *t* = -2.54, *p* = .011).

Accuracy[19]. Neither the main effect of Prime Type (β = 0.41, *SE* = 0.34, *z* = 1.23, *p* = .219) nor that of Prime Condition (β = 0.13, *SE* = 0.11, *z* = 1.16, *p* = .247) was significant. The interaction between Prime Type and Prime Condition (β = -0.02, *SE* = 0.11, *z* = -0.18, *p* = .855) was also non-significant.

Whether form facilitation was significant under each prime condition was also inspected. As shown in Table 17, priming was significant in none of the prime conditions.

---

[18] Model formula (lmer): RT_Inverse ~ Related * Condition + (1 | Participant) + (1 | Item)
[19] Model formula (lmer): Acc ~ Related * Condition + (1 | Participant) + (1 | Item)

Table 16

Three L2 Speaker Groups' Mean Reaction Times (in Milliseconds) and Error Rates (in Parentheses) under Different Prime Conditions in the Masked Priming Lexical Decision Task

	Homophone		/l/ - /ɹ/		/p/ - /f/		/d/ - /t/ or /b/ - /t/		Orthographic neighbors sharing the initial		Nonwords	
	Related (*meet-MEAT*)	Unrelated (*kids-MEAT*)	Related (*read-LEAD*)	Unrelated (*wish-LEAD*)	Related (*full-PULL*)	Unrelated (*sick-PULL*)	Related (*dear-TEAR*)	Unrelated (*boys-TEAR*)	Related (*play-PLAN*)	Unrelated (*best-PLAN*)	Related (*name-MAME*)	Unrelated (*long-MAME*)
Low-proficiency L2 speakers (*n* = 30)	804.7 (10.7%)	837.3 (12.4%)	789.2 (12.5%)	812.4 (13.6%)	807.2 (7.1%)	828.7 (8.5%)	778.6 (3.2%)	799.7 (10.0%)	841.8 (7.8%)	800.5 (8.5%)	899.5 (10.5%)	903.8 (8.0%)
	32.6* (1.7%)		**23.2*** (1.1%)		21.5† (1.4%)		21.1 **(6.8%)***		-41.3† (0.7%)		4.3 **(-2.5%)***	
Medium-proficiency L2 speakers (*n* = 30)	718.5 (9.8%)	745.9 (15.1%)	740.8 (11.9%)	768.3 (14.0%)	734.9 (6.1%)	748.4 (7.8%)	734.1 (8.6%)	722.3 (11.0%)	750.5 (6.4%)	734.0 (9.3%)	797.4 (7.2%)	797.3 (7.0%)
	27.4* (5.3%)†		**27.5*** (2.0%)		13.5 (1.7%)		-11.8 (2.4%)		-16.5 (2.9%)		-0.1 (-0.2%)	
High-proficiency L2 speakers (*n* = 30)	814.4 (2.5%)	819.6 (5.6%)	834.7 (7.9%)	845.7 (7.3%)	834.3 (5.7%)	838.6 (3.0%)	775 (2.5%)	790.9 (3.8%)	830.6 (4.7%)	828.7 (3.8%)	873.4 (3.8%)	886.2 (3.6%)
	5.2 (3.1%)†		11.0 (-0.6%)		4.3 (-2.7%)		15.9 (1.3%)		-1.9 (-0.9%)		12.8 (-0.2 %)	
L1 speakers (*n* = 30)	657.0 (5.1%)	640.2 (4.7%)	685.4 (9.7%)	665.8 (5.6%)	671.7 (6.8%)	655.2 (5.2%)	657 (4.7%)	649.2 (3.8%)	654.1 (6.4%)	655.8 (4.2%)	700.8 (4.6%)	704.1 (4.4%)
	-16.8 (-0.4%)		-19.6† (-4.2%)†		-16.5 (-1.6%)		-7.8 (-0.9%)		1.7 (-2.2%)		3.3 (-0.2%)	

Notes. Significance codes: '**' < .01, '*' < .05, †: $p < .10$

Table 17

Form Priming under Five Prime Conditions Observed in Each of Three L2 Speaker Groups

Group	Prime Condition	Reaction time			Error rate		
		Priming	*t*	*p*	Priming	*z*	*p*
Low-prof. L2 speakers	Homophone	32.6 ms	-2.61	.009**	1.7%	0.63	.527
	/l/-/ɹ/	23.2 ms	-2.27	.023*	1.1%	0.46	.644
	/f/-/p/	21.5 ms	-1.66	.096†	1.4%	0.72	.471
	/b/-/t/ or /d/-/t/	21.1 ms	-0.75	.456	6.8%	2.63	.008**
	Pairs sharing the initial	-41.3 ms	1.73	.084†	0.7%	0.08	.938
Medium-prof. L2 speakers	Homophone	27.4 ms	-2.34	.020*	5.3%	1.90	.057†
	/l/-/ɹ/	27.5 ms	-2.54	.011*	2.0%	0.55	.583
	/f/-/p/	13.5 ms	-1.09	.278	1.7%	0.69	.490
	/b/-/t/ or /d/-/t/	-11.8 ms	0.14	.886	2.4%	0.90	.366
	Pairs sharing the initial	-16.5 ms	1.10	.272	2.9%	1.24	.215
High-prof. L2 speakers	Homophone	5.2 ms	-0.70	.483	3.1%	1.70	.090†
	/l/-/ɹ/	11.0 ms	-1.37	.170	-0.6%	-0.20	.846
	/f/-/p/	4.3 ms	0.16	.874	-2.7%	-1.55	.122
	/b/-/t/ or /d/-/t/	15.9 ms	-1.12	.262	1.3%	0.78	.436
	Pairs sharing the initial	-1.9 ms	0.28	.783	-0.9%	-0.59	.554

Notes. Significance codes: '**' < .01, '*' < .05, †: $p < .10$

4.3.3. High-proficiency L2 Speakers

Response Latency[20]**.** Table 16 shows that despite their higher proficiency, the high-proficiency L2 speakers' reaction times were overall slower than of the other two L2 speaker groups (see Figure 1). This is thought to be because the participants in this group were significantly older (M = 35.0) than those in the low- (M = 24.0) and medium-proficiency groups (M = 25.5).

In the high-proficiency group, neither the main effect of Prime condition (β = -0.17, SE = 8.15, t = -0.02, p = .983) nor that of Prime Type (β = -29.22, SE = 25.16, t = -1.16, p = .246) was significant. The interaction between Prime Type and Prime Condition was also non-significant (β = 5.27, SE = 7.55, t = 0.70, p = .485).

Whether form facilitation was significant under each prime condition was also inspected. As shown in Table 17, priming was significant in none of the prime conditions.

Accuracy[21]**.** Neither the main effect of Prime Type (β = 0.49, SE = 0.47, z = 1.03, p = .304) nor that of Prime Condition (β = 0.12, SE = 0.12, z = 1.00, p = .319) was significant. The interaction between Prime Type and Prime Condition (β = -0.17, SE = 0.15, z = -1.149, p = .251) was also non-significant.

Whether form facilitation was significant under each prime condition was also inspected. As shown in Table 17, priming was significant in none of the prime conditions.

4.3.4. Comparison between L1 Speakers and High-Proficiency L2 Speakers

One of the predictions this study had made before conducting the experiment was that high-proficiency L2 speakers would show a nativelike pattern of form priming. Therefore, to

[20] Model formula (lmer): RT_Inverse ~ Related * Condition + (1 | Participant) + (1 | Item)
[21] Model formula (lmer): Acc ~ Related * Condition + (1 | Participant) + (1 | Item)

determine if this prediction would be supported, L1 and high-proficiency L2 speakers' correct reaction times to LDT word targets were compared. For statistical analysis, the same predictors and model construction procedures that are described in Section 4.2 were used, except that instead of the whole L2 speaker data set, only high-proficiency L2 speaker data were used.

Table 18

Output from a Linear-Mixed Effects Model for Inverse-Transformed Reaction Times in L1 and High-proficiency L2 Speakers' Masked Priming Lexical Decision

Fixed effects				
Predictor	Estimate	*SE*	*t*	*p*
(1) The reference category for Group: L1 speakers				
(Intercept)	-1595.65	48.08	-33.21	< .001***
Related	38.53	28.92	1.33	.183
Condition	-2.72	6.94	-0.39	.695
GroupL2	275.54	62.25	4.43	< .001***
Related:Condition	-4.73	8.71	-0.54	.587
Related:GroupL2	-66.38	35.35	-1.88	.060
Condition:GroupL2	2.57	5.31	0.48	.629
Related:Condition:GroupL2	9.87	10.62	0.93	.353
(2) The reference category for Group: L2 speakers				
(Intercept)	-1321.11	48.08	-27.48	< .001***
Related	-27.85	28.93	-0.96	.336
Condition	-0.16	6.93	-0.02	.982
Related:Condition	5.14	8.69	0.59	.555
Random effects	Variance	SD	Corr	
Participant	53441	231.17		
Item	5440	73.76		
Related	3080	55.50	0.08	

Notes. Formula (lmer): RT_Inverse ~ Related * Condition * Group + (1 | Participant) + (1 + Related | Item), Significance codes: '***' < .001, '**' < .01, '*' < .05

The outputs of the model are shown in Table 18. The most important result is that when low- and medium-proficiency L2 speakers' data were excluded, the interaction between Prime

Type and Group was non-significant (β = -66.38, *SE* = 35.35, *t* = -1.88, *p* = .060) although it approached significance. In addition, when the reference category for Group was switched to high-proficiency L2 speakers, as mentioned earlier, neither the main effect of Prime Type (β = -0.16, *SE* = 6.93, *t* = -0.02, *p* = .982) nor the interaction between Prime Type and Prime Condition (β = 5.14, *SE* = 8.69, *t* = 0.59, *p* = .555) was significant. In brief, high-proficiency L2 speakers overall showed a nativelike pattern of form priming (a null form priming effect), and the strength of form facilitation was not different across the prime conditions.

4.3.5. The Effect of Proficiency on Form Priming in the Entire Sample of L2 Speakers

Another prediction of this study was that as L2 speakers' proficiency improves, they will show weaker form facilitation. To test this prediction, how proficiency affected form priming was examined using the entire L2 speaker data set. A linear mixed-effects model was used for reaction time analysis, and a logistic mixed-effects model was used for accuracy analysis. In addition to Prime Type and Prime Condition, standardized cloze test scores (proficiency) were fitted as a fixed effect in these models. The same model building procedure described above was adopted.

The outputs of the models are presented in Table 19. Even though Figure 4 shows a trend that correlates higher L2 speaker proficiency with weaker form facilitation both in reaction time and accuracy, the results of statistical analysis reveal that the effect of proficiency on form priming was non-significant. In both models, only the main effect of proficiency on correct response time was significant (β = -50.96, *SE* = 21.13, *t* = -2.41, *p* = .016) indicating that the reaction times of more proficient L2 speakers were significantly faster. The main effects of proficiency on accuracy and its interactions with other predictors were all non-significant.

Table 19

Output from Mixed-Effects Models Exploring the Effect of Proficiency on Form Priming for L2 Speakers' Reaction Time and Accuracy

Fixed effects	Reaction Times				Accuracy			
Predictor	Estimate	*SE*	*t*	*p*	Estimate	*SE*	*z*	*p*
(Intercept)	-1354	32.78	-41.31	< .001***	2.51	0.32	7.84	< .001***
Related	-77.28	18.84	-4.10	< .001***	0.32	0.21	1.51	.131
Condition	-0.01	7.88	-0.01	.991	0.17	0.10	1.83	.068
ClozeZ	-50.96	21.13	-2.41	.016*	0.19	0.15	1.29	.196
Related:Condition	18.12	5.64	3.21	.001**	-0.03	0.07	-0.42	.675
Related:ClozeZ	22.46	15.16	1.48	.139	-0.12	0.22	-0.56	.579
Condition:ClozeZ	-0.78	2.25	-0.35	.727	-0.03	0.04	-0.85	.398
Related:Condition:ClozeZ	-6.37	4.53	-1.41	.159	0.03	0.07	0.38	.703
Random effects	Variance	SD	Corr		Variance	SD	Corr	
Participant	35077	187.29			0.38	0.62		
Item	9109	95.44			1.19	1.09		
Related	1822	42.69	-0.30					
ClozeZ					0.05	0.23	-1.00	

Notes. Formula for reaction times (lmer): RT_Inverse ~ Related * Condition * SpellZ * VocabZ + (1 | Participant) + (1 + Related | Item), Formula for accuracy (glmer): Acc ~ Related * Condition * ClozeZ + (1 | Participant) + (1 + ClozeZ | Item), Significance codes: '***' < .001, '**' < .01, '*' < .05.

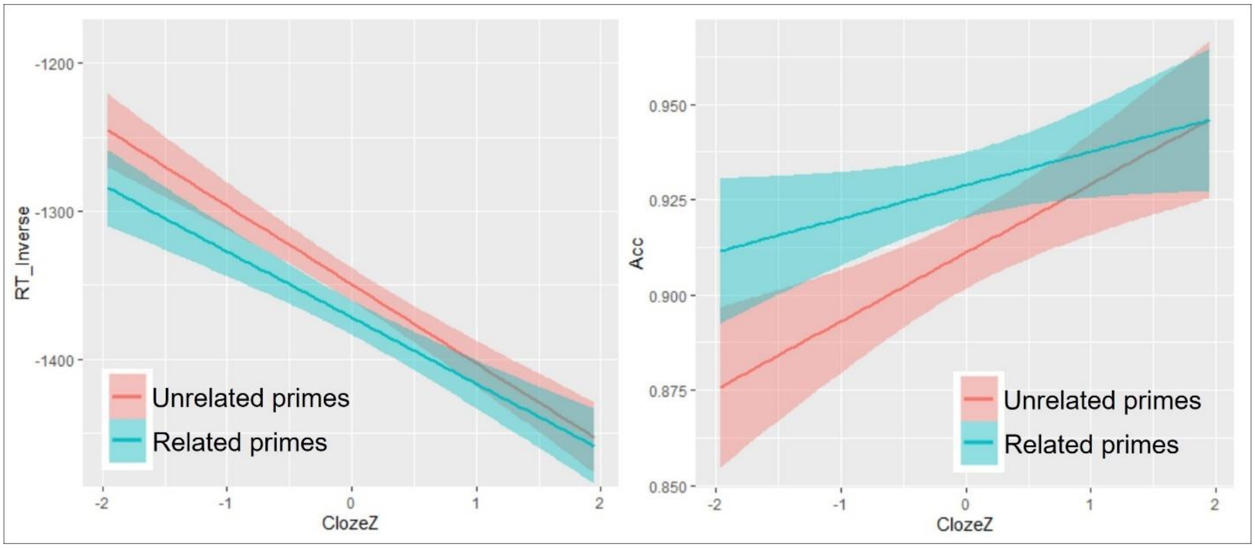

Figure 4. The Interaction between Proficiency and Prime Type for L2 Speakers' Reaction time (Left) and Accuracy (Right).

4.3.6. Summary

Form facilitation was overall significant in low- and medium-proficiency L2 speakers. On the other hand, high-proficiency L2 speakers showed a null priming effect like L1 speakers. These overall patterns of priming are consistent with the FLR hypothesis which assumes the developmental nature of orthographic representations as a function of L2 experience. In addition, the finding that medium-proficiency L2 speakers who did not show facilitation for *dear-TEAR* still showed facilitation for *read-LEAD* was expected.

However, an unpredicted finding was also obtained. That is, even low-proficiency L2 speakers did not show significant facilitation in reaction time for *dear-TEAR* although the facilitation in accuracy for this pair was significant. Form facilitation in accuracy for L2 speakers was observed by Nakayama and Lupker (2018), but they also detected significant facilitation in reaction time at the same time. Thus, because previous research has usually focused on facilitation in reaction time, it is difficult to conclude that facilitation in accuracy alone supports our original hypothesis for prime-target pairs without a confusing contrast in low-proficiency L2 speakers. Furthermore, low-proficiency L2 speakers did not show significant facilitation under the /f/-/p/ contrast condition either. If it is supposed that the /f/-/p/ contrast might not have posed a serious difficulty for accurate phonological encoding of words having this contrast (as shown in Section 4.1), the stimuli used under this condition (e.g., *full-PULL*) should be processed similarly to the prime-target pairs without a nonnative contrast (e.g., *dear-TEAR*). Therefore, the finding that neither significant facilitation in accuracy nor in reaction time was observed for *full-PULL* suggests that it may be problematic to consider the facilitation in accuracy for *dear-TEAR* very meaningful. In short, no strong evidence for orthographic form facilitation for minimal pairs without a nonnative contrast was obtained.

The previous section (Section 4.3.5) also reported the effect of global proficiency on orthographic neighbor priming using this predictor as a continuous variable in a statistical model. The results presented above indicate that global proficiency did not significantly modulate form priming. Nevertheless, it should be noted that what the FLR hypothesis actually supposes is that more fully-specified orthographic representations (orthographic precision) rather than proficiency will modulate orthographic neighbor priming. The reason why the predictions were made based on global proficiency in this paper was just due to a simple assumption that the higher a student's proficiency was, the more precise their orthographic representations would be. However, previous research shows that L2 proficiency does not necessarily correlate with orthographic precision (Park, 2021), and orthographic precision has been considered a separate construct distinguished from reading skills (Andrews et al., 2020; Andrews & Hersch, 2010; Andrews & Lo, 2012; Frith, 1985). Therefore, in the next section, the effect of orthographic precision, the key construct for the FLR hypothesis, on orthographic form priming is analyzed and reported. In addition, the effects of other theoretically motivated constructs (vocabulary size and accurate sound perception of nonnative contrasts) on orthographic form priming are inspected.

4.4. The Relationships between Proficiency, Orthographic Precision, Vocabulary Size and Sound Perception Ability

Before exploring the effects of orthographic precision, vocabulary size and sound perception ability on form orthographic neighbor priming, correlations analyses were first performed with the mean scores that each L2 participant had obtained in all the measures to see how these constructs were related to one another. Two-tailed Pearson correlation analyses were used when two variables that were compared satisfied the assumptions of linearity and normality. When one of two variables did not meet these assumptions, non-parametric tests (two-tailed Spearman's rho) were employed. L1 speakers' data were not included in this analysis.

The results are summarized in Table 20. This table presents significant correlations between all measures except between vocabulary size and identification of words with the /f/-/p/ contrast, $r(88) = .16$, $p = .126$. Of course, interpreting this table requires caution considering a high Type I error rate which was caused by multiple comparisons. When a Bonferroni-adjusted alpha value was used (0.002) taking into account the number of comparisons shown in Table 18 (0.05 ÷ 21 times), some of the correlations between identification of the /f/-/p/ contrast and the outputs of other measures were not significant.[22] However, the table at least clearly shows that

[22] Specifically, the correlations between /f/-/p/ identification in words and proficiency ($r(88) = .29$, $p = .005$), between /f/-/p/ identification in nonwords and proficiency ($r(88) = .23$, $p = .032$), between /f/-/p/ identification in nonwords and vocabulary size ($r(88) = .27$, $p = .010$), between /f/-/p/ identification in nonwords and orthographic

the higher participants' proficiency, the more precise (i.e., the spelling scores) and bigger (i.e., the vocabular scores) their mental lexicon.

Figure 5 illustrates the relationship between proficiency and sound identification ability. It shows that compared to perception of contrasts that are also distinguished in L2 participants' L1 (the /b/-/t/ and /d/-/t-/ contrasts), it was more difficult for L2 speakers to correctly perceive nonnative contrasts (the /f/-/p/ and /l/-/ɹ/ contrasts). Furthermore, it demonstrates that the /f/-/p/ contrast was easier for them to learn, such that it was acquired at a relatively earlier developmental stage than the /l/-/ɹ/ contrast. The figure also suggests that the reason for the weaker correlation between the outputs of written measures and identification of the /f/-/p/ contrast was because even beginner L2 speakers were already quite accurate in differentiating the phoneme /f/ from /p/. Finally, Figure 5 shows that although the most advanced L2 learners were able to distinguish the /l/-/ɹ/ contrast as accurately as the other contrasts when stimuli were words, they still could not differentiate this contrast very accurately if it was embedded in nonwords.

To sum up, improvements in orthographic precision and vocabulary size by a function of proficiency were observed. It was also found that participants were able to perceive nonnative phonemes more accurately as their proficiency increased. The following sections will examine how the development of each of the theory-driven constructs affected form priming in a masked orthographic neighbor priming LDT.

---

precision ($r(88) = .31$, $p = .003$) and between /f/-/p/ identification in nonwords and /f/-/p/identification in words ($r(88) = .31$, $p = .003$) were non-significant with the adjusted alpha (0.002).

4.5. The Effects of Orthographic Precision and Vocabulary Size on Orthographic Priming

Because no significant main effect of Prime Type or its interactions with other predictors was found in accuracy or reaction time analyses for nonword targets, subsequent analyses focus on participants' correct response latencies to word targets.

To investigate whether form priming was modulated by orthographic precision and vocabulary size, L1 and L2 speakers' data were examined separately. Inverse-transformed participants' correct responses to targets were the dependent variable. To explore which construct played a more significant role in form priming between orthographic precision and vocabulary size, both SpellZ (standardized spelling scores) and VocabZ (standardized vocabulary scores), in addition to Prime Type and Prime Condition, were fitted at the same time as fixed-effects, and Participant and Item were fitted as random effects. The same model building procedure that was described in Section 4.1 was adopted.

Table 20

Bilateral Correlations between Proficiency, Vocabulary Size, Orthographic Precision, /f/-/p/ and /l/-/ɹ/ Identification in Words, and /f/-/p/ and /l/-/ɹ/ Identification in Nonwords

		1 (*n* = 90)	2 (*n* = 90)	3 (*n* = 90)	4 (*n* = 90)	5 (*n* = 90)	6 (*n* = 90)	7 (*n* = 90)
1. Cloze	Pearson Correlation	1						
	p	-						
2. Vocabulary	Pearson Correlation	.591**	1					
	p	< .001	-					
3. Spelling	Pearson Correlation	.579**	.432**	1				
	p	< .001	< .001	-				
4. /f/-/p/ identification in words	Spearman's rho	.292**	.163	.319**	1			
	p	.005	.126	.002	-			
5. /l/-/ɹ/ identification in words	Spearman's rho	.432**	.377**	.364**	.333**	1		
	p	< .001	< .001	< .001	.001	-		
7. /f/-/p/ identification in nonwords	Spearman's rho	.226*	.271**	.311**	.494**	.461**	1	
	p	.032	.010	003	< .001	< .001	-	
8. /l/-/ɹ/ identification in words	Spearman's rho	.384**	.327**	.361**	.306**	.864**	.413**	1
	p	< .001	.002	.001	.003	< .001	< .001	-

Notes. Significance codes: '**' < .01, '*' < .05

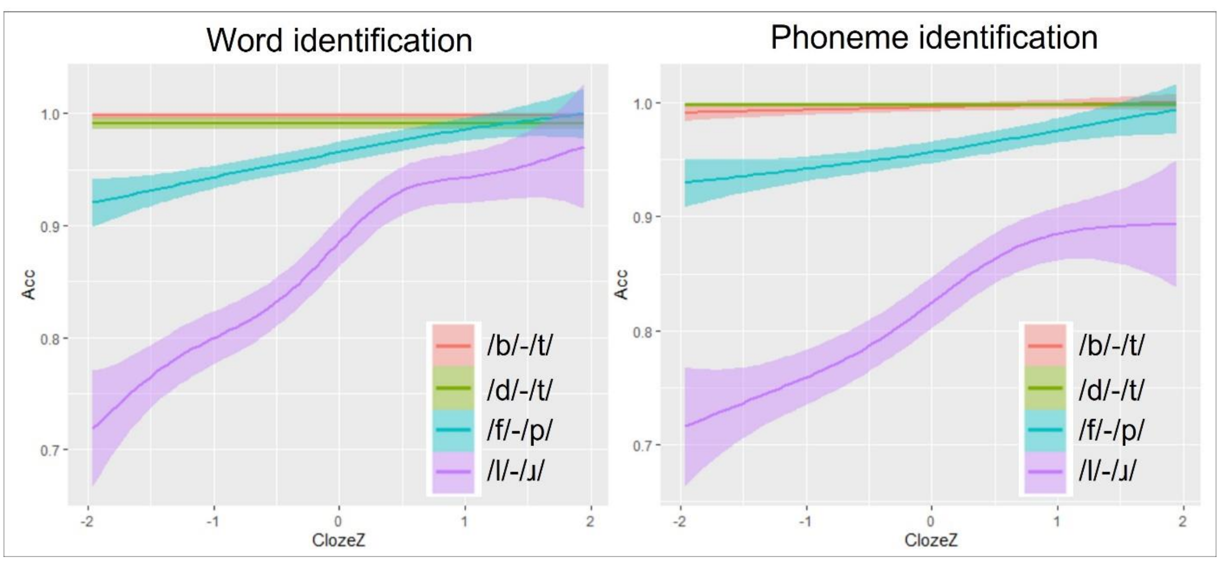

Figure 5. The Relationships between Proficiency and Word Identification (Left) and between Proficiency and Phoneme Identification (Right).

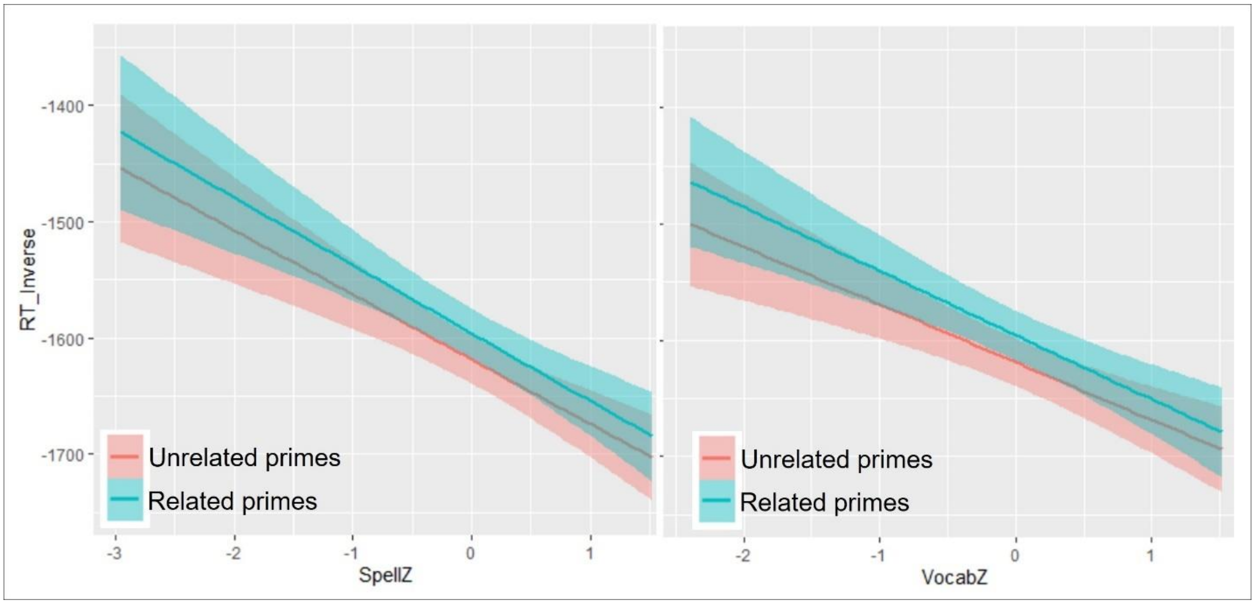

Figure 6. The Interactions between Prime Type and Orthographic Precision (Left) and between Prime Type and Vocabulary Size (Right) in L1 speakers.

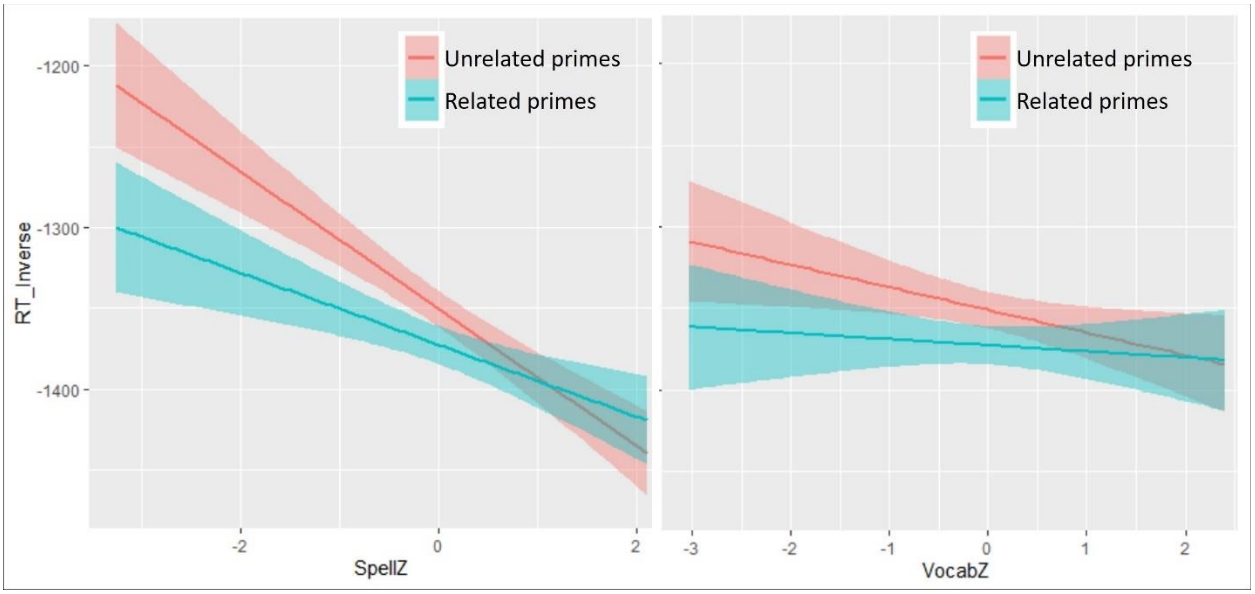

Figure 7. The Interactions between Prime Type and Orthographic Precision (Left) and between Prime Type and Vocabulary Size (Right) in L2 speakers.

Table 21

Output from Models Exploring the Effects of Orthographic Precision and Vocabulary Size in L1 and L2 Speakers

Fixed effects	L1 speakers				L2 speakers			
Predictor	Estimate	*SE*	*t*	*p*	Estimate	*SE*	*t*	*p*
(Intercept)	-1597.11	64.62	-24.71	< .001***	-1358.69	34.13	-39.82	< .001***
Related	25.53	36.98	0.69	.490	-83.54	19.72	-4.24	< .001***
Condition	-5.36	6.50	-0.83	.409	0.73	7.90	0.09	.926
SpellZ	-51.54	62.41	-0.83	.409	-27.43	24.56	-1.12	.264
VocabZ	1.92	63.07	0.03	.976	5.85	24.65	0.24	.813
Related:Condition	-4.27	11.11	-0.38	.701	20.30	5.90	3.44	< .001***
Related:SpellZ	19.09	31.87	0.60	.549	52.21	17.25	3.02	.002**
Condition:SpellZ	3.21	4.71	0.68	.495	-2.77	2.55	-1.09	.276
Related:VocabZ	-11.00	31.68	-0.35	.729	14.28	17.15	0.83	.405
Condition:VocabZ	-9.44	4.72	-2.00	.046*	-0.51	2.55	-0.20	.841
SpellZ:VocabZ	-0.69	64.12	-0.01	.991	10.38	19.40	0.54	.593
Related:Condition:SpellZ	-5.62	9.50	-0.59	.554	-13.73	5.14	-2.67	.008**
Related:Condition:VocabZ	3.39	9.46	0.36	.720	-3.44	5.09	-0.68	.499
Related:SpellZ:VocabZ	21.25	33.29	0.64	.523	12.44	13.37	0.93	.352
Condition:SpellZ:VocabZ	4.35	4.90	0.89	.374	-1.74	2.00	-0.87	.384
Related:Condition:SpellZ:VocabZ	-0.52	9.99	-0.05	.956	-4.62	4.00	-1.15	.249
Random effects	Variance	SD	Corr		Variance	SD	Corr	
Participant	69981	264.54			30547	193.77		
Item	3117	55.83			9041	95.03		
Related	4918	70.13	0.38		1834	42.82	-0.30	

Notes. Formula (lmer): RT_Inverse ~ Related * Condition * SpellZ * VocabZ + (1 | Participant) + (1 + Related | Item),
Significance codes: ‘***’ < .001, ‘**’ < .01, ‘*’ < .05

4.5.1. L1 Speakers

The outputs of models are shown in Table 21. Neither the main effect of SpellZ (β = -51.54, *SE* = 62.41, *t* = -0.83, *p* = .409) nor VocabZ (β = 1.92, *SE* = 63.07, *t* = 0.03, *p* = .976) was significant. As illustrated in Figure 6, the interactions between Prime Type and SpellZ (β = 19.09, *SE* = 31.87, *t* = 0.60, *p* = .549) and between Prime Type and VocabZ (β = -11.00, *SE* = 31.68, *t* = -0.35, *p* = .729) were also non-significant. Unexpectedly, a significant interaction between Prime Condition and VocabZ was found (β = -9.44, *SE* = 4.72, *t* = -2.00, *p* = .046). To understand the nature of this interaction, Prime Condition was refitted as a factor variable. The main effect of VocabZ under each of the five different prime conditions was then examined changing the reference category. The results reveal that there was a slight difference in the strength of the main effect of VocabZ across the homophone (β = -12.43, *SE* = 62.82, *t* = -0.20, *p* = .843), the /l/-/ɹ/ contrast (β = -6.65, *SE* = 62.98, *t* = -0.11, *p* = .916), the /f/-/p/ contrast (β = -33.44, *SE* = 62.87, *t* = -0.53, *p* = .595), the /b/-/t/ or /b/-t/ contrast (β = -31.42, *SE* = 62.77, *t* = -0.50, *p* = .616) and the same initial (β = -46.51, *SE* = 62.78, *t* = -0.74, *p* = .459) conditions. However, under none of the conditions was the main effect of VocabZ significant.

4.5.2. L2 Speakers

The results of analysis for L2 speakers are also presented in Table 21. The main effects of SpellZ (β = -27.43, *SE* = 24.56, *t* = -1.12, *p* = .264) and VocabZ (β = 5.85, *SE* = 24.65, *t* = 0.24, *p* = .813) were not significant. However, although the interaction between Prime Type and VocabZ was non-significant (β = 14.28, *SE* = 17.15, *t* = 0.83, *p* = .405), a significant interaction between Prime Type and SpellZ was detected (β = 52.21, *SE* = 17.25, *t* = 3.02, *p* = .002) indicating that the more precise L2 speakers' orthographic representations, the weaker the form facilitation in

L2 speakers (see Figure 7). Moreover, a significant three-way interaction between Prime Type, Prime Condition and SpellZ was observed (β = -13.73, *SE* = 5.14, *t* = -2.67, *p* = .008) suggesting that the strength of the interaction between Prime Type and SpellZ differed across the five prime conditions.

Table 22

The Interaction between Prime Type and Orthographic Precision under Five Prime Conditions

Prime Condition	Estimate	*SE*	*t*	*p*
Homophone	24.96	16.59	1.51	.132
/l/-/ɹ/	46.39	16.56	2.80	.005**
/f/-/p/	-7.64	16.29	-0.47	.639
/b/-/t/ or /d/-/t/	23.14	16.16	1.43	.152
Pairs sharing the initial	-31.11	15.86	-1.96	.049*

As shown in Table 22 and Figure 8, SpellZ significantly modulated the form priming under the /l/-/ɹ/ contrast condition (β = 46.39, *SE* = 16.56, *t* = 2.80, *p* = .005). Figure 8 seems to suggest that SpellZ also substantially influenced the form priming for homophone pairs, even though the interaction between Prime Type and SpellZ did not reach significance when the effect of VocabZ was controlled for (β = 24.96, *SE* = 16.59, *t* = 1.51, *p* = .132). In addition, a significant interaction between Prime Type and SpellZ was observed for prime-target pairs sharing the initial letter (β = -31.11, *SE* = 15.86, *t* = -1.96, *p* = .049). However, interestingly, for these pairs, the direction of the interaction was reverse. In other words, the more precise L2 speakers' orthographic representations, the weaker the form inhibition.

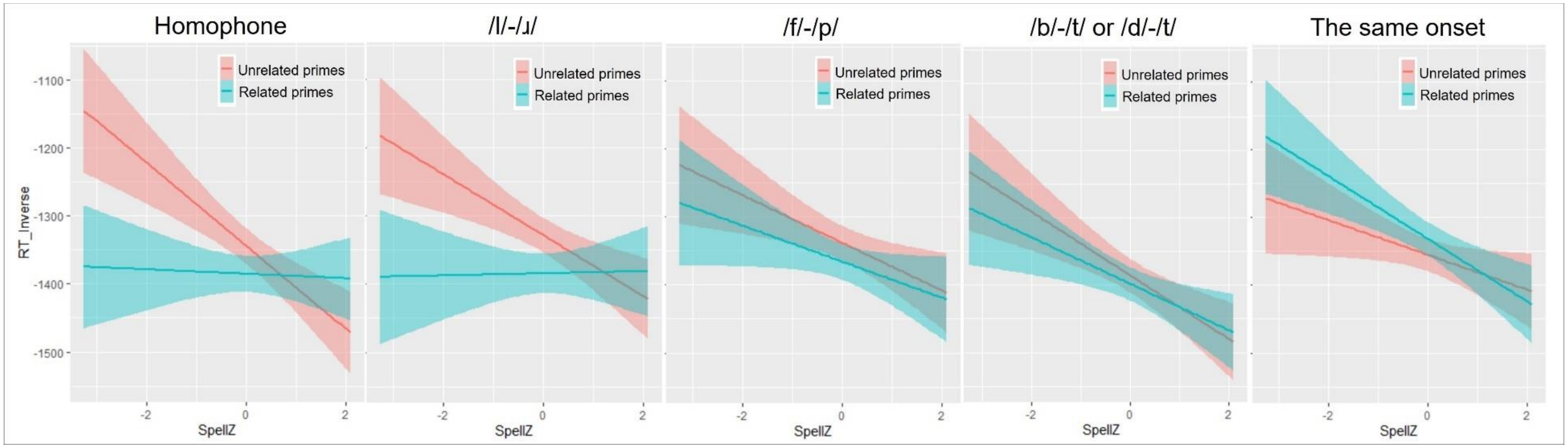

Figure 8. The interactions between Prime Type and SpellZ under Five Different Prime Conditions in L2 Speakers.

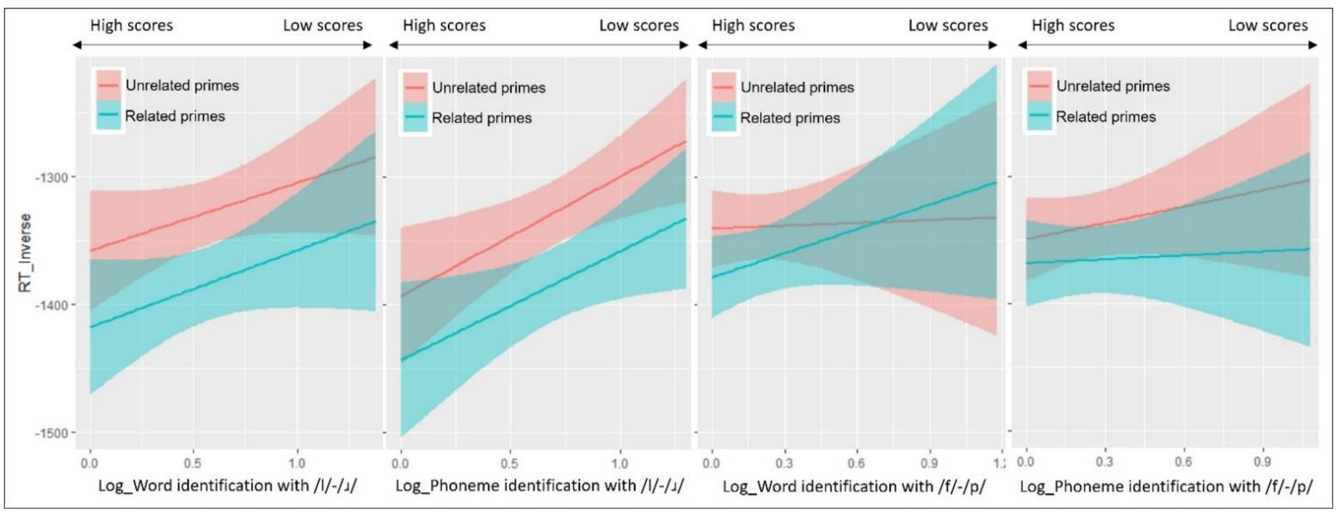

Figure 9. Interactions between Prime Type and Word or Phoneme Identification Scores under the /l/-/ɹ/ (Left) and /f/-/p/ (Right) Contrast Conditions in L2 Speakers (Note. Word and phoneme identification scores were log-transformed after reflection, which means a distance-preserved transformation through the X axis due to negative skewness. For that reason, the left side of each figure represents high scores whereas the right side represents low scores).

In addition to the finding that the effect of priming for prime-target pairs with a shared initial letter (e.g., *play-PLAN*) was 17 ms of inhibition, although this was non-significant (β = -24.93, *SE* = 18.36, *t* = 1.36, *p* = .175) (see Table 16), the finding that orthographic precision modulated the magnitude of form inhibition rather than facilitation under the same initial condition is difficult to understand. Potential reasons for these findings are discussed in Section 5.3.

4.6. The Relationship between Sound Perception and Orthographic Priming

As noted earlier, it was expected that if L2 speakers had precise phonological representations for nonnative phonological contrasts, it would be evident in their sound perception ability. This is because if they were able to accurately perceive contrasts, the accurate perception might help them establish more precise phonological representations when exposed to aural L2 input (e.g., Pallier et al., 2001). At the same time, representations were also expected to become the reason for better perception because fine-grained phonological representations would assist L2 speakers' accurate categorization of L2 phonological segments (e.g., Flege, 2003; Meador et al., 2000; Thorn & Gathercole, 2001). Therefore, since we guessed that the facilitation in L2 speakers for prime-target pairs with the /l/-/ɹ/ contrast detected in the LDT might originate from these participants' fuzzy phonological representations, we predicted that good performers in the word identification and phoneme identification tasks would show weaker form facilitation in the LDT than poorer L2 sound perceivers.

To test this hypothesis, after reflection (*reflection* means a statistical method of distance-preserved transformation through the X axis using the following formula: the maximum score of the task +1 - each participant's score), we first log-transformed participants' scores for the /l/-/ɹ/

and the /f/-/p/ contrasts in the word and phoneme identification tasks because those scores were negatively skewed. Then, VocabZ was dropped from the model that was used for examining the effects of orthographic precision and vocabulary size (described in the previous section) because the effect of this variable was minimal (see Table 21). After that, to examine whether the correct sound identification of words with the /l/-/ɹ/ contrast affected form facilitation in the LDT, word identification scores for this phonological contrast were fitted as a fixed factor in the model in which the reference category for Prime Condition was the condition with the /l/-/ɹ/ contrast. To clarify, in this model, the dependent variable was the inverse-transformed L2 speakers' correct reaction times to word targets, and the predictors were Prime Type, Prime Condition, SpellZ and word identification scores for the /l/-/ɹ/ contrast. Random effects were Participant and Item, and a random slope for Prime Type was added to Item. The model output for the interaction between Prime Type and word identification scores was then inspected. The result shows that this interaction was non-significant (β = 58.28, *SE* = 43.05, *t* = 1.35, *p* = .176) (see also Figure 9 above). The same analysis was performed using another model in which word identification scores were replaced with phoneme identification scores because word identification could have been affected by the lexicality of stimuli that might interfere with measuring pure sound perception ability. The result of the analysis reveals that the interaction between phoneme identification scores and Prime Type was not significant either (β = 57.40, *SE* = 42.01, *t* = 1.37, *p* = .172).

For an analysis of the /f/-/p/ contrast, after switching the reference category for Prime Condition to the condition with the /f/-/p/ contrast, word identification scores for the /f/-/p/ contrast were fitted as a fixed factor. The interaction between Prime Type and word identification scores was not significant (β = 44.24, *SE* = 62.31, *t* = 0.71, *p* = .478) (see also Figure 9). The output of another model in which the word identification scores for the /f/-/p/ contrast were replaced with phoneme identification scores for the same contrast also shows that the interaction between Prime Type and phoneme identification scores was non-significant (β = -4.15, *SE* = 52.63, *t* = -0.08, *p* = .937).

Although the interactions between Prime Type and the identification of the two phonological contrasts were non-significant, this could have been affected by the high skewness of the data. In other words, participants' summed word or phoneme identification scores were highly skewed even after transformation, so they were not very suitable for the statistical models described above. Therefore, additional analysis was performed to make sure that L2 speakers' sound perception ability indeed did not influence the form facilitation found in the LDT. In this analysis, instead of summed scores, the correctness of each participant's response to each word identification stimulus was used as a predictor. Note that in the word identification task, the same stimuli that were used as primes and targets in the LDT were reused. We reasoned that if a participant had a precise phonological representation of a particular word such as *lead*, he/she would have been able to accurately identify both *read* and *lead* during the word identification task. We then considered that the impact of correct identification of *read* and *lead* on the form facilitation for a particular prime-target pair (*read-LEAD*) might be observable.

For this analysis, participants' responses in the word identification task were recoded. If a participant correctly identified both the prime (e.g., *read*) and the target (e.g., *LEAD*) in the auditory perception task, it was coded as '2'. If he/she correctly identified only either the prime or the target, it was coded as '1'. If he/she was incorrect both for the prime and the target, it was coded as '0'. Each participant's individual score for each stimulus was then used as a predictor (WordID) in the mixed-effects model. After that, the reference category for Prime Condition was set as the condition with the /l/-/ɹ/ contrast,[23] and the interaction between Prime Type and WordID was examined. The result suggests that this interaction was non-significant (β = 20.80, *SE* = 32.42, *t* = 0.64, *p* = .520). The same procedure was followed after switching the reference category for Prime Condition to the /f/-/p/ contrast condition. The interaction between Prime Type and WordID was not significant either (β = -1.70, *SE* = 68.73, *t* = -0.03, *p* = .980).

Finally, to determine whether form facilitation for pairs with the /l/-/ɹ/ contrast was non-significant if L2 speakers received a perfect score for this contrast in the word identification task, an additional analysis was performed. For this, out of L2 speakers' whole data set, 13 L2 participants who correctly identified all the stimuli with /l/-/ɹ/ contrast were coded as a single group ('0') while the other L2 participants were coded as another group ('1'). This group variable was fitted as a predictor (a factor variable) instead of WordID into a model in which the reference category for Prime Condition was set as the /l/-/ɹ/ contrast condition[24] to determine the main effect of Prime Type. This perfect word identification group (*n* = 13) showed 33.3 ms of facilitation. The result reveals that this facilitation was significant (β = -112.17, *SE* = 46.58, *t* = -2.41, *p* = .016), indicating that even L2 speakers who could correctly identify all the words with the /l/-/ɹ/ contrast still showed significant form facilitation. The other group who did not obtain a perfect score in the word identification task (*n* = 77) also showed 18.3 ms of significant facilitation (β = -51.73, *SE* = 19.45, *t* = -2.66, *p* = .008).

---

[23] The formula: RT_Inverse ~ Related [Prime Type] * Condition2 [Prime Condition with the /l/-/ɹ/ contrast as the reference category] * SpellZ * WordID [the correctness of each participant's word identification responses to the prime and target] + (1 | Participant) + (1 + Related | Item)

[24] The formula: RT_Inverse ~ Related [Prime Type] * Condition2 [Prime Condition with the /l/-/ɹ/ contrast as the reference category] * SpellZ * GroupID [the reference category for the L2 speaker group who received the perfect identification score for words the /l/-/ɹ/ contrast] + (1 | Participant) + (1 + Related | Item)

The same analysis was performed using the phoneme identification scores. The perfect phoneme identification group (n = 8) showed 30.5 ms of facilitation. Although this facilitation was non-significant (β = -85.95, *SE* = 65.85, *t* = -1.31, *p* = .192), it was thought to be because of the small sample size of this group. The other group who did not obtain a perfect score in the word identification task (*n* = 82) showed 19.5 ms of significant facilitation (β = -48.51, *SE* = 18.82, *t* = -2.58, *p* = .010).

In short, no evidence was found for the hypothesis that L2 speakers' word or phoneme identification of a nonnative contrast would modulate the strength of form facilitation for prime-target pairs with this contrast.

4.7. Performance of L2 Speakers with Precise Orthographic Representations in Large Mental Lexicon

In Section, 4.3.5, even though both L1 and high-proficiency L2 speakers showed a null priming effect, Table 16 illustrates that whereas L1 speakers showed a trend toward inhibition, high-proficiency L2 speakers still showed a trend toward facilitation. Based on the assumption that L1 and L2 lexicons are not fundamentally distinct, and on the alleged role of vocabulary size in visual word recognition (e.g., Brysbaert et al., 2017), this difference is suspected to be because of the difference in vocabulary size. That is, as demonstrated in Figure 10, while the quality of orthographic representations of high-proficiency L2 speakers was comparable to that of L1 speakers, their vocabulary size was still significantly smaller. Thus, we performed further analysis to explore whether form inhibition would be observable even in the L2 mental lexicon if it became comparable to the L1 lexicon in terms of both level of precision and size.

For this analysis, we summed participants' spelling and vocabulary scores after changing them into percentage scores (the number of correct responses ÷ the total number of items in each task). The results are presented in Figure 11. They show that L1 speakers' mental lexicon is superior to L2 speakers' lexicon in consideration of both quality and size. Even though high-proficiency L2 speakers' average summed score was lower than L1 speakers' average, there were five exceptional L2 speakers (two who were provided with List 1 and three who were provided with List 2 during the LDT) who obtained a score that was higher than L1 speakers' average Vocabulary + Spelling score. We examined their reaction times in the LDT. For counterbalancing of the lists, one L2 speaker (who was provided with List 1 in the LDT) whose score was almost the same as L1 speakers' average score was also included. Table 23 shows the reaction times of these six L2 speakers.

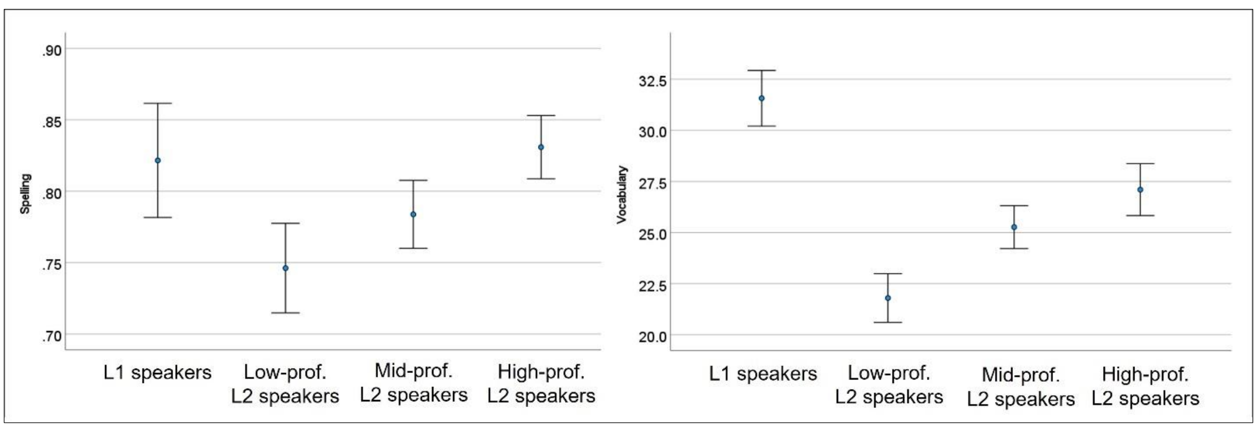

Figure 10. Levels of Orthographic Precision (left) and Vocabulary Size (right) of an L1 and Three L2 Speaker Groups.

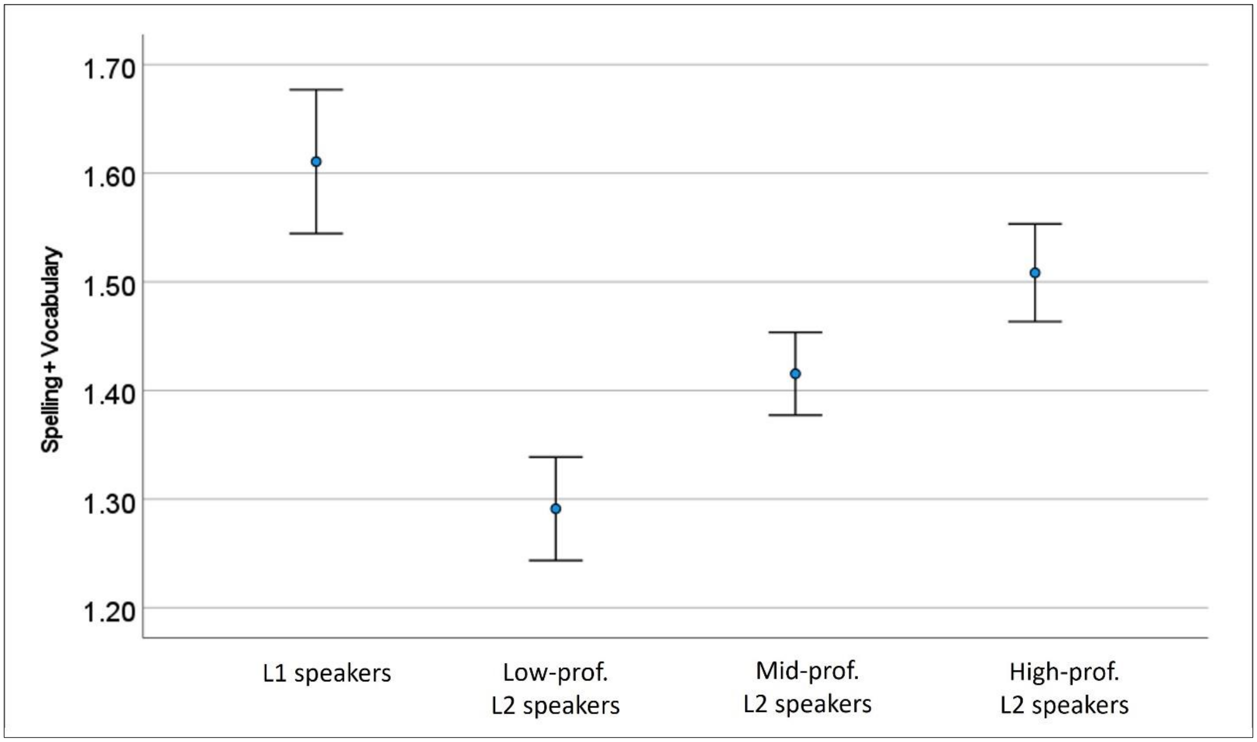

Figure 11. An L1 and Three L2 Speaker Groups' Summed Scores of Spelling and Vocabulary Tasks.

Table 23

Mean Reaction Times (in Milliseconds) of Six L2 Speakers in Masked Priming Lexical Decision Who Performed Better than average L1 Speakers in the Spelling and Vocabulary Tasks

Homophone		/l/ - /ɹ/		/p/ - /f/		/d/ - /t/ or /b/ - /t		The same initial	
Related (*meet-MEAT*)	Unrelated (*kids-MEET*)	Related (*read-LEAD*)	Unrelated (*wish-LEAD*)	Related (*full-PULL*)	Unrelated (*sick-PULL*)	Related (*down-TOWN)*	Unrelated (*make-TOWN)*	Related (*play-PLAN*)	Unrelated (*best-PLAN*)
834.8	809.1	864.0	765.8	827.1	842.9	774.5	742.4	825.1	840.0
-25.8		-98.2		15.8		-32.1		14.9	

As shown in Table 23, these six outstanding L2 speakers showed an overall trend toward inhibition (-24.4 ms). For statistical analysis, these six participants were fitted as a separate group in the model described in Section 4.3.5, and the main effect of Prime Type was then examined after setting this new group as the reference category for Group. The results indicate that inhibition was not significant in this group (β = 82.96, *SE* = 57.37, *t* = 1.45, *p* = .148), probably due to the small sample size. Nevertheless, Table 23 suggests that form inhibition may be possible even in L2 speakers if they establish a sufficiently precise and large L2 mental lexicon.

4.8. The Relationship between the Effect of Vocabulary Size and the Length of Words

Unlike our prediction that vocabulary size as well as orthographic precision would influence form priming, as described in Section 4.5 this effect was not detected in either L1 or L2 speakers (see Table 21 and Figure 7). Considering that L1 speakers' vocabulary size was much larger than L2 speakers' (see Figure 10 and Figure 11), the null effect of vocabulary size on form priming in L1 speakers was suspected to be because all L1 speakers, even the lowest scorers in the vocabulary task, had a mental lexicon that was large (and precise) enough to produce form inhibition (see the next chapter for more discussion). However, the null effect of vocabulary size for L2 speakers was not consistent with previous research. In Park's study (2021), the strength of form facilitation in L2 speakers was modulated by proficiency, and it was interpreted that proficiency might be related to vocabulary size.

The masked priming experiments used in both Park's study (2021) and the present study were conducted under the same environment. However, a few differences can be found in the stimuli. The first difference is that whereas prime-target pairs had no phonological relationship in Park's study (2021), many homophone and near-homophone pairs were used in the current study. Therefore, if it is assumed that masked facilitative phonological priming indeed occurred for L2 speakers in the current study, because vocabulary size has nothing to do with pre-activation of phonemes, its modulating role in the form facilitation might not be detected. Another difference is that the present study used shorter stimuli compared to those in Park's study. Specifically, while only four- and five-letter words were used in Park's study (2021), in the current study almost a third of the stimuli (31.3%) were three-letter words. Previous research suggests that vocabulary size contributes to developing precise orthographic representations. However, if orthographic precision is accounted for through the spelling task as in the current study, vocabulary size would be more closely related to neighborhood size, which has been considered an important factor triggering lexical competition (Andrews & Hersch, 2010; Davis & Lupker, 2006; Forster et al., 1987; Nakayama et al., 2008). If so, regarding the argument that certain amount of overlap is required for form inhibition (Dufour & Peereman, 2003b; Gor & Cook, 2020; Slowiaczek & Hamburger, 1992), it is possible to assume that three-letter words would not be sufficient to activate the neighborhood. If the neighborhood of short words was not activated due to little overlap, it would be the reason for the weak impact of vocabulary size on lexical competition. For that reason, we examined whether the effect of vocabulary size on form priming was weaker for short words.

For this analysis, Length (of stimuli) was added as an additional fixed factor into the model that was used for an earlier analysis for determining the effects of SpellZ and VocabZ in L2 speakers (see Section 4.4). This variable was fitted as a factor variable to observe a possible interaction between VocabZ and Prime Type under a different reference category for Length.

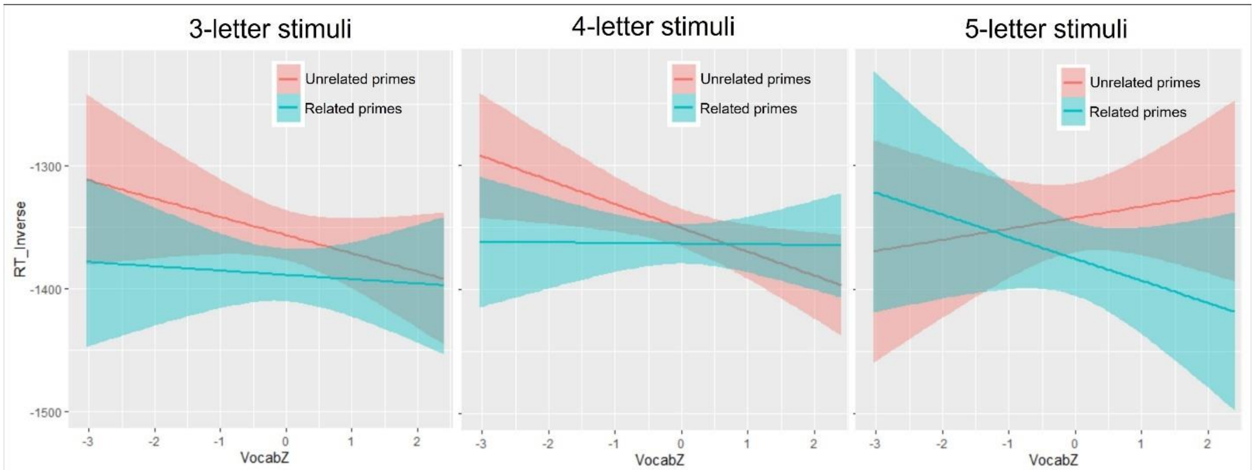

Figure 12. The Interactions between Prime Type and VocabZ with Three-letter, Four-letter and Five-letter Stimuli.

The interaction between Prime Type and VocabZ was significant when four-letter stimuli (which consisted of 55% of the word targets) were used (β = 48.18, *SE* = 23.41, *t* = 2.06, *p* = .040) (see Figure 12). However, this interaction was non-significant when three-letter (31.3% of the word targets) (β = -10.52, *SE* = 31.09, *t* = -0.34, *p* = .735) and five-letter stimuli (13.7%) were used (β = -79.89, *SE* = 49.57, *t* = 1.61, *p* = .107). The non-significant effect of vocabulary size on form priming with three-letter stimuli is consistent with our prediction. However, the non-significant interaction between Prime Type and VocabZ with five-letter stimuli was unexpected. Subsequent investigation determined that the five-letter stimuli had lower neighborhood density (*M* = 6.3) compared to that of the three- (*M* = 19.3) or four-letter stimuli (*M* = 14.1). It is worth noting that previous studies that reported the impact of neighborhood density on form priming (Andrews & Hersch, 2010; Davis & Lupker, 2006; Forster et al., 1987; Nakayama et al., 2008) always used stimuli with a wide range of density (e.g., *M* = 3.8 for low vs. *M* = 15.2 for high density words in Nakayama et al, 2008). Therefore, based on the assumption that the interaction between VocabZ and form priming indicates that when vocabulary size (neighborhood density) is large, lexical competition is strong, it can be reasoned that the five-letter stimuli might have an insufficient number of neighbors such that the impact of vocabulary size on form priming for them was not detected.

To sum up, the findings of the investigation on the influence of word length suggest that a potential reason for the small impact of vocabulary size on form priming that was observed in the current study might be minimal form overlap (in the case of three-letter words) and low neighborhood density of the stimuli (in the case of five-letter words). These characteristics of the stimuli might have created an environment that was not conducive for detecting the effect of vocabulary size.

4.9. Analysis of L2 Speakers' Translation Responses

In Section 2.2.3, three plausible mechanisms of form facilitation for prime-target pairs with the difficult phonological contrast were suggested. The first reason is the situation in which L2 speakers have trouble orthographically encoding L2 words that have a phoneme for which transfer of a relevant L1 phoneme is not straightforward (Wang & Geva, 2003). Put differently, it was thought that if L2 speakers are not sure which phoneme should be mapped onto an L2 grapheme, orthographic encoding would be challenging for them. If so, form facilitation would be produced by weak orthographic representations operationalized as the failure in successful mapping of orthography to any meaning.

We considered that the initial challenge for orthographic encoding might be observable in L2 participants' familiarity ratings on LDT stimuli. In the familiarity rating task, three options were provided for each stimulus: (a) I have never seen this word before, (b) I do not know the meaning of this word and (c) I know the meaning of this word. Thus, if a participant chose the

Familiarity Rating (B), it would mean that he/she remembered having seen the stimulus but had not yet succeeded in form-meaning mapping. Therefore, we reasoned that if the inclusion of a nonnative phoneme in the word form poses an extra challenge for orthographic encoding, the Rating (B) would be more frequently selected for words having this phoneme. Note that the stimuli used under the five prime conditions of the LDT were comparable in frequency. For that reason, although word frequency for L2 speakers could be different from that for L1 speakers, it was thought possible that any difference in the number of responses with the Rating (B) across the five prime conditions could be at least partially attributed to whether the stimuli had a nonnative phoneme.

Before the analysis, all familiarity rating responses for words that were used as unrelated primes in the LDT were excluded from participants' familiarity rating data since they were not informative. To give an example for clarification, the unrelated prime for the target *LEAD* was *wish* whereas its related prime was *read*. Thus, the unrelated prime did not include a nonnative phoneme such as /l/ or /ɹ/. Therefore, since *wish* did not represent stimuli for the /l/-/ɹ/ contrast condition, it was excluded. Out of the remaining data set with 10,800 familiarity rating responses, 133 responses with the "I have never seen this word before" rating (1.2%) were also excluded.

Descriptive statistics for this analysis are summarized in Table 24. For statistical analysis, a mixed-effect logistic regression model was used. In the model, participants' familiarity ratings were the dependent variable ("I do not know the meaning of this word" coded as '1' and "I know the meaning of this word" coded as '0'). Participant Group (low- (coded as '1') vs. mid- (coded as '2' vs. high-proficiency L2 speakers (coded as '3')) and stimulus Condition (the /l/-/ɹ/ contrast condition coded as '1', the /f/-/p/ contrast condition coded as '2', the /b/-/t/ or /d/-/t/ contrast condition coded as '3', the homophone condition coded as '4' and the condition with the same initial coded as '5') were fitted as fixed effects. Participant and Item were random effects. The same model construction procedure used for earlier analyses was followed.

Table 24.

The Numbers of "I Do Not Know the Meaning of This Word" Responses in the Familiarity Rating Task across Three L2 Participant Groups and across Five Stimulus Conditions in the Familiarity Rating Task

Participant	Homophone (4.9%)	/l/-/ɹ/ (9.8%)	/f/-/p/ (6.7%)	/b/-/t/ or /d/-/t/ (5.0%)	Sharing the initial (2.0%)
Low-prof. L2 speakers	50 (7.3%)	113 (16.1%)	86 (12.3%)	54 (7.8%)	26 (3.6%)
Mid-prof. L2 speakers	44 (6.2%)	69 (9.6%)	49 (6.9%)	44 (6.2%)	17 (2.4%)
High-prof. L2 speakers	11 (1.5%)	27 (3.8%)	8 (1.1%)	9 (1.3%)	0 (0%)

The results of the statistical analysis show that there was a significant difference in the probability of selecting the Familiarity Rating (B) across the three L2 speaker groups (β = -1.35, SE = 0.23, z = -5.90, p < .001) (see Appendix J for specific model outputs). Specifically speaking, it was found that the more proficient the participants were, the less frequently they selected the Rating (B). These results indicate that as participants' proficiency increased, the probability of unsuccessful orthographic encoding decreased. At the same time, the results reveal a significant main effect of Condition (β = -0.64, SE = 0.21, z = -3.00, p = .002) indicating that the probability of unsuccessful initial orthographic encoding differed across the stimulus conditions. On the other hand, no significant interaction between Group and Condition was found (β = -0.04, SE = 0.06, z = -0.67, p = .506).

Subsequent analysis was performed using the same model, after changing the variable Condition to a factor variable and setting the condition with the /b/-/t/ or /d/-/t/ contrast (the control condition) as the reference category. The results reveal the difference in the probability of choosing the “I do not know the meaning of the word” option between the /l/-/r/ contrast condition (9.8%) and the /b/-/t/ or /d/-/t/ contrast condition (5.0%) was significant ($\beta = 1.93$, $SE = 0.94$, $z = 2.05$, $p = .040$). The difference between the /f/-/p/ contrast (6.7%) and the /b/-/t/ or /d/-/t/ contrast condition (5.0%) also approached significance ($\beta = 1.83$, $SE = 0.96$, $z = 1.92$, $p = .055$). (See Appendix J for the whole model outputs.) These results suggest that L2 speakers might find it more difficult to encode the spellings of L2 words if these words include a nonnative phoneme. Of course, in the current study, other factors that might influence vocabulary acquisition, such as semantic concreteness, cognateness (e.g., De Groot & Keijzer, 2000) or contextual or semantic diversity (e.g., Hamrick & Pandža, 2020), were not controlled. Therefore, great caution is required for interpretation. However, the results are in line with the possibility that nonnative phonemes may exert negative influence on orthographic encoding at the initial learning stage. In addition, since the LDT responses with the Familiarity Rating (B) were included for earlier analyses, it could be considered that because there were more word stimuli with very weak orthographic representations (even without mapping onto any meanings) under the /l/-/ɹ/ and the /f/-p/ contrast conditions, these weak orthographic representations might have had some influence on form priming for prime-target pairs with these two contrasts.

Another way to gauge the levels of fuzziness in orthographic representations was to look at the consistency between familiarity ratings and accuracy of responses in the translation task. In other words, it was expected that if orthographic representations of stimuli that were used under a certain prime condition were indeed particularly fuzzier than those used under other prime conditions, more incorrect translation responses would be observed for stimuli that were used under the former prime condition. This is because if participants retained coarse-grained orthographic representations, the form-meaning mappings even for words that they felt familiar with would be imprecise. Therefore, a set of incorrect L1 translation equivalents for L2 words that were perceived to be familiar was examined.

The descriptive statistics of this analysis are summarized in Table 25. Overall, phonologically related prime-target pairs (i.e., the pairs under the homophone and the near-homophone conditions) show higher error rates. Table 26 illustrates that across the stimulus conditions, many of unsuccessful orthographic form-meaning mappings were caused by phonology. For instance, for the stimuli *site* and *lap*, L2 speakers sometimes provided L1 translations of *cite* and *wrap*. These responses indicate that similar or identical phonological forms of other words interfered with accurate spelling-meaning mappings for stimulus words. Subsequent analysis discovered that the reason for more frequent incorrect L1 translations shown in Table 25 was because there were more words that had confusing phonological forms under the homophone, the /l/-/ɹ/ contrast and the /f/-/p/ contrast conditions compared to the other stimulus conditions. When inspecting inaccurate orthography-semantic mappings that were not caused by the same or similar phonological forms, the probability of incorrect L1 translation did not greatly vary across the five stimulus conditions as shown in Table 27.

A statistical analysis was performed to ensure that the L2 speakers indeed more frequently provided incorrect L1 translation equivalents for words (which they thought they knew) from the homophone and the near-homophone conditions. For the analysis, out of the words that were indicated to be known in the familiarity rating task (10,071 responses), correctly translated words were coded as ‘1’, and incorrectly translated words were coded as ‘0’.

Participant Group and stimulus Condition (the /l/-/ɹ/ contrast condition coded as ‘1’, the homophone condition coded as ‘2’, the /f/-/p/ contrast condition coded as ‘3’, the /b/-/t/ or /d/-/t/ contrast condition coded as ‘4’, and the condition with the same initial coded as ‘5’) were fitted as fixed effects. Participant and Item were random effects. The same model construction procedure used for earlier analyses was followed.

Table 25.

Probability of Providing an Incorrect L1 Translation Equivalent for an L2 Word That was Indicated to Be Known

Participant	Homophone (9.4%)	/l/-/ɹ/ (10.4%)	/f/-/p/ (7.2%)	/b/-/t/ or /d/-/t/ (3.8%)	Sharing the initial (2.6%)
Low-prof. L2 speakers	14.4%	16.0%	12.4%	5.1%	4.3%
Mid-prof. L2 speakers	9.1%	10.7%	7.7%	4.9%	2.3%
High-prof. L2 speakers	5.1%	5.5%	3.0%	1.5%	1.3%

The results show that the main effect of Group (β = 0.93, *SE* = 0.16, z = 5.71, p < .001), as well as the main effect of Condition (β = 0.61, *SE* = 0.16, z = 3.85, p < .001) was significant (see Appendix K for full model outputs). A subsequent analysis also reveals that the probability of inaccurate spelling-meaning mapping was significantly higher for words with the /l/-/ɹ/ contrast (β = -1.80, *SE* = 0.70, z = -2.57, p = .010), the /f/-/p/ contrast (β = -1.57, *SE* = 0.70, z = -2.23, p = .026) and homophones (β = -1.56, *SE* = 0.70, z = -2.23, p = .026) than those without a confusing contrast (i.e., words with the /b/-/t/ or /d/-/t/ contrast). This means that there were more words with unreliable orthography-meaning mappings under the former three conditions compared to the latter condition.

Another suspected reason for form facilitation was the repetition priming caused by poorly encoded orthographic representations. Specifically, it was reasoned that if what was activated by the prime during the LDT was actually the target, it would produce facilitation. To explore this possibility, the cases of translation responses where participants provided a translation equivalent of a prime (e.g., *read*) for a target (e.g., *lead*) or vice versa were examined. Through this investigation, it was found that 4.5% of the responses under the /l/-/ɹ/ contrast condition and 3.0% of the responses under the /f/-/p/ contrast, as well as 5.1% of the responses under the homophone condition, resulted from form confusion between the prime and the target. This confusion of the prime with the target, or vice versa, was rarely observed for words with the /b/-/t/ or /d/-/t/ contrast (0.04%). These findings also suggest that homophones and near-homophones posed an extra challenge for orthographic encoding.

To conclude, the results of the analyses of L2 speakers' familiarity ratings and translation responses suggest that L2 speakers are likely to have poorer orthographic representations if word stimuli included a nonnative phoneme or shared the same pronunciation (or similar pronunciations) with another word. This suggests that the significant facilitation observed under the homophone and the /l/-/ɹ/ contrast conditions in the LDT might be greatly influenced by fuzzy orthographic representations. Of course, the possibility that the facilitation might come from the pre-activation of the target phonemes by the prime cannot be totally disregarded. In the present studies, 80.8% of the responses to LDT stimuli with the /l/-/ɹ/ stimuli and 86.0% of the responses to LDT stimuli with /f/-/p/ contrast were indicated to be known and correctly translated. These data may indicate that the orthographic representations of many stimuli were not that fuzzy. However, Park (2021) reported that just being able to provide a correct L1 translation equivalent for a stimulus does not guarantee lexical competition. In his study, even some stimuli that were perceived to be familiar and correctly translated still produced form facilitation unless they were high-frequency words or when L2 speakers' orthographic representations lacked precision. All these findings suggest that the possibility should not be underestimated that the locus of form facilitation for homophones and minimal pairs with the /l/-/ɹ/ contrast is weak orthographic representations of these words.

Table 26.

Proportion of Phonology-Based Incorrect Form-Meaning Mappings for L2 Words That Was Indicated to Be Known

<table>
<tr><th rowspan="2">Participant</th><th colspan="2">Homophone</th><th colspan="2">/l/-/ɹ/</th><th colspan="2">/f/-/p/</th><th colspan="2">/b/-/t/ or /d/-/t/</th><th colspan="2">Sharing the initial</th></tr>
<tr><th>Homophone (55.1%)</th><th>Near-homophone (19.6%)</th><th>Homophone (5.5%)</th><th>Near-homophone (49.3%)</th><th>Homophone (3.4%)</th><th>Near-homophone (45.3%)</th><th>Homophone (14.3%)</th><th>Near-homophone (7.8%)</th><th>Homophone (50.9%)</th><th>Near-homophone (3.6%)</th></tr>
<tr><td>Examples</td><td>meet-meat</td><td>heel-hill</td><td>rap-wrap</td><td>load-road</td><td>full-fool</td><td>fin-pin</td><td>dear-deer</td><td>dip-deep</td><td>tale-tail</td><td>bat-bet</td></tr>
<tr><td>Low-prof. L2 speakers</td><td>52.2%</td><td>15.2%</td><td>5.3%</td><td>56.4%</td><td>5.3%</td><td>57.9%</td><td>9.1%</td><td>12.1%</td><td>53.3%</td><td>6.7%</td></tr>
<tr><td>Mid-prof. L2 speakers</td><td>59.0%</td><td>22.6%</td><td>8.7%</td><td>40.6%</td><td>2.0%</td><td>33.3%</td><td>18.2%</td><td>6.1%</td><td>31.2%</td><td>0.0%</td></tr>
<tr><td>High-prof. L2 speakers</td><td>55.6%</td><td>25.0%</td><td>0.0%</td><td>47.4%</td><td>0.0%</td><td>28.6%</td><td>18.2%</td><td>0.0%</td><td>85.7%</td><td>0.0%</td></tr>
</table>

Table 27.

Probability of Providing an Incorrect L1 Translation Equivalent for an L2 Word That was Indicated to Be Known after Accounting for Phonology-Based Error

Participant	Homophone (2.4%)	/l/-/ɹ/ (4.7%)	/f/-/p/ (3.8%)	/b/-/t/ or /d/-/t/ (3.0%)	Sharing the initial (1.2%)
Low-prof. L2 speakers	4.7%	6.1%	4.6%	4.0%	1.7%
Mid-prof. L2 speakers	1.6%	5.4%	5.0%	3.7%	1.6%
High-prof. L2 speakers	1.0%	2.9%	2.1%	1.3%	0.2%

Chapter 5. Discussion

The present study was conducted to track the developmental trajectory of L2 orthographic representations within the framework of the FLR hypothesis. This study first replicated the finding of previous research (Jiang, 2021; Nakayama & Lupker, 2018; Park, 2021; Qiao & Forster, 2017) that L1 and L2 speakers show different patterns of form priming. In the current study, whereas L1 speakers showed a null priming effect, L2 speakers showed significant form facilitation. In addition, it was observed that the strength of form priming was different across prime conditions in L2 speakers while it was not in L1 speakers. In the L2, form facilitation was stronger for homophone pairs (e.g., *meet-MEAT*) and prime-target pairs with the /l/-/ɹ/ contrast (e.g., *read-LEAD*), but it was not significant if the prime and the target did not share a difficult phonological contrast (e.g., *dear-TEAR*). Although the /f/-/p/ contrast is also a nonnative phonological contrast, L2 participants in the present study showed no significant form facilitation for minimal pairs with this contrast either. The difference in the strengths of form priming across different priming conditions that was observed in L2 speakers cannot be explained by L2 speakers' slower or unsuccessful processing of the prime since all participants were tested under the same five prime conditions and since the stimuli had comparable linguistic characteristics across the conditions. Therefore, these differences suggest that what determined the strength of form facilitation under each prime condition was the quality of the encoding of lexical representations.

Subsequently employed individual-differences measures show that the strength of the form facilitation observed in L2 speakers was modulated by orthographic precision. Even though the weak facilitation observed for prime-target pairs with the /b/-/t/ or /d/-/t/ contrast, albeit at an earlier developmental stage, was different from our prediction, the finding that orthographic precision significantly modulated the facilitative orthographic neighbor priming effect strongly supports the FLR hypothesis. At the same time, the finding that vocabulary scores modulated the form priming for four-letter words is also in the same line with the alleged important role of vocabulary size in L2 lexical development (e.g., Brysbaert et al., 2017; Diadone & Darcy, 2021; Llompart, 2021).

The results of the word and the phoneme identification tasks indicate that L2 speakers were less accurate in identification of the /l/-/ɹ/ contrast compared to L1 speakers. The /f/-/p/ contrast was also perceived less accurately when it was embedded in nonwords. This finding suggests that although it was relatively easier for L2 speakers to acquire the phonological categories of /f/ and /p/, accurate perception of this contrast was not as straightforward as perception of L2 contrasts that are also discriminated in their L1 (the /b/-/t/ and /d/-/t/ contrast). Even though the current study observed improvement in sound identification ability for the /l/-/ɹ/ and the /f/-/p/ contrasts as a function of proficiency, no evidence was found for the assumption that better sound perception would produce weaker form facilitation for minimal pairs with these nonnative contrasts.

This chapter discusses the implications of these findings and several questions these findings bring up.

5.1. The Impact of a Nonnative Phonological Contrast on Visual Word Recognition

The significant form facilitation for prime-target pairs with the /l/-/ɹ/ contrast observed in the LDT is considered to have originated from the fuzzy form (orthographic or phonological) representations of L2 words with this contrast. L2 speakers showed that they were significantly less accurate in perceiving this contrast in the word and the phoneme identification tasks compared to L1 speakers. Because L2 speakers had not been able to perceive the difference between /l/ and /ɹ/ in L2 aural input as inferred from the results, they would not have been able to develop precise phonological representations of this contrast. In addition, the inaccurate phonological representations could have posed a challenge for the establishment of accurate orthographic representations.

Three possible specific mechanisms to explain how the fuzzy phonological representations affect form priming in visual word identification were suggested in Section 2.2.3. The first was the situation where L2 speakers find it difficult to map an L2 grapheme onto a nonnative L2 phoneme since they are not sure about which phoneme should be mapped onto this grapheme (Wang & Geva, 2003). L2 participants in the current study gave the “I do not know the meaning of this word” rating more frequently for words with the /l/-/ɹ/ contrast than for other words without a nonnative contrast (the /b/-/t/ and /d/-/t/ contrast) (see Section 3. 8). Thus, this finding supports the plausibility of the first mechanism.

The second mechanism was form facilitation from fuzzy orthographic representations. In the translation task, L2 speakers sometimes provided an L1 translation equivalent of the prime for the target or vice versa if they had the /l/-/ɹ/ contrast. They also sometimes produced incorrect L1 translations for words whose phonological forms were confusable with another word. These findings support the plausibility of the second mechanism. However, more than 80% of the LDT stimuli were perceived to be familiar and translated correctly in the familiarity rating and translation task. This finding indicates that the third mechanism might have played a role as well. This mechanism supposes that since the phonological form that was activated by the prime was not clearly distinguishable from that of the target, it might have provided a phonological boost for the target just as homophone pairs did in the present study.

If the third mechanism indeed operated, it offers an additional research question. Are the two phonological forms of a nonnative contrast represented in L2 speakers' mental lexicon (e.g., /l/ in *lead* and /ɹ/ in *read*) identical since they are both mapped onto a single phoneme that has correspondences with its L1 counterpart (e.g., "ㄹ") (Pallier et al., 2001)? Or, are the two forms mapped onto two distinct phonemes but are not discriminable from each other because the phonological categories of these two phonemes are fuzzy? When we examined the data from L2 speakers, the facilitation in the /l/-/ɹ/ direction was significant (40 ms) whereas it was non-significant in the /ɹ/-/l/ direction (13 ms). Similarly, under the /f/-/p/ contrast condition, the facilitation in the /p/-/f/ direction was significant (54 ms) whereas it was not in the /f/-/p/ direction (3 ms).

When Schmidt (1995) asked native Korean-speaking participants in her study to give a similarity rating between each English phoneme and an L1 phoneme that sounded similar to the L2 phoneme, Korean participants gave a similarity rating 3.90 for /l/-"ㄹ", 2.97 for /ɹ/-"ㄹ", 4.54 for /p/-"ㅍ" and 2.21 for /f/-"ㅍ". Considering Schmidt's findings (1995), it seems that form facilitation might occur only when the prime had a phoneme similar to a relevant L1 phoneme. This is an example of asymmetry in form facilitation. A similar asymmetry was reported in a previous auditory phonological priming study (Barrios et al., 2016; see also Weber & Cutler, 2004 that reports asymmetry in the L2 observed in experiments using the visual world paradigm). If this asymmetry in form priming indeed occurred in the present study, it suggests the establishment of two separate representations for each phoneme of a nonnative contrast. However, because we did not control the linguistic characteristics of stimuli (e.g., length or word frequency or neighborhood density) before comparing the strength of facilitation in different prime directions, we do not make a strong claim regarding this possibility.

Finally, it is worth noting how phonological representations of words with a nonnative contrast, which are not very difficult to discriminate, are represented in the mental lexicon. In the present study, L2 speakers showed relatively good perception of the /f/-/p/ contrast. To be more specific, they were as accurate as L1 speakers in identification of the /f/-/p/ contrast at least if this contrast was embedded in words. Consequently, they did not show significant form facilitation for minimal pairs with this contrast in the LDT. This may suggest that L2 speakers were able to develop precise phonological representations of words having this contrast at a relatively earlier developmental stage than those with the /l/-/ɹ/ contrast. However, it is not yet determined whether the phonological representations of words with /p/ and /f/ were equally robust. As noted in the previous paragraph, in our data set, L2 speakers appeared to show significant facilitation if the priming direction was from a word with /p/ to another with /f/. However, out of 16 prime-target pairs used under the /f/-/p/ contrast condition, only 3 pairs were presented in this direction. This was because a high-frequency word in each minimal pair was used as the prime while the other member was used as the target. Among the selected 16 pairs, only three had words that started with /p/ and had a higher frequency than the other member of the pair. Therefore, if more words starting with /p/ had been used as the prime, significant facilitation might have been observed as was for the minimal pairs used under the /l/-/ɹ/ contrast. Future research is necessary to clearly understand the exact nature of L2 phonological representations with nonnative contrasts.

5.2. The Orthographic and Phonological Priming in Visual L2 Word Recognition

The current study predicted that at an earlier stage of lexical development, a strong facilitative orthographic priming effect would be observed because of fuzzy orthographic representations. However, unlike in previous studies (Jiang, 2021; Nakayama & Lupker, 2018; Qiao & Forster, 2017), even when L2 speakers' proficiency was low, or when the quality of their orthographic representations was poor (see Figure 7), form facilitation was not very strong unless the prime and target were homophones (e.g., *meet*-*MEAT*) or near-homophones (e.g., *read*-*LEAD*). In other words, even though L2 speakers showed significant form facilitation overall, this facilitation appears to have mainly been driven by homophone and near-homophone pairs. On the other hand, as noted in Section 4.3, none of the three L2 proficiency groups showed significant facilitation under the /f/-/p/ contrast, the /b/-/t/ or /d/-/t/ contrast, and the same initial conditions.

A reason for this may be the characteristics of the stimuli used in the present study. Unlike in the previous studies (Jiang, 2021; Nakayama & Lupker, 2018; Qiao & Forster, 2017) in which stimuli with longer than four letters were used, the current experiment used many three-letter stimuli (31.3%), and the mean length of stimuli (3.8) was also shorter than those used in the other studies (4 ~ 7 letters in Jiang, 2021; M = 4.4 or 4.6 in Nakayama & Lupker, 2018; 6 ~ 8 letters in Qiao & Forster, 2017). When short stimuli are used, there is less overlap between the prime and the target compared to when long stimuli are used. Therefore, for short stimuli there would be less facilitative boost produced by overlap in sub-lexical components (letters) between the prime and the target (Elgort, 2011; Lupker & Davis, 2006). This may explain why orthographic form facilitation was weak even in the low-proficiency L2 speakers, in contrast to our expectation.

Contrary to the weak orthographic priming for words without a difficult phonological contrast, form facilitation was stronger in L2 speakers if prime-target pairs had a nonnative phonological contrast. As discussed in the previous section, there are two interpretations regarding the locus this facilitation. The first interpretation assumes that the significant facilitation for such pairs was produced by fuzzy orthographic representations. When several words have the same pronunciation, it will interrupt the development of precise orthographic representations. In the present study, L2 speakers frequently showed incorrect form-meaning mappings for words whose pronunciations were not clearly distinguishable from those of other words (i.e., homophone and near-homophone pairs). This discovery indicates that the orthographic representations of homophones or near-homophones were fuzzier than those of words without a difficult contrast. In this interpretation, the reason why L1 speakers did not show significant facilitation for homophones should be because they had precise orthographic representations of L1 homophones.

A second interpretation is based on the alleged facilitative masked phonological priming that occurred during visual lexical access. This facilitation, produced by an overlap in phonological codes (phonemes), has been observed in both within- and between-language priming experiments (see Sections 2.2.2 and 2.2.4). Previous research on the time course of activation of orthographic and phonological information of the form prime (Ferrand & Grainger, 1992; Perfetti & Tan, 1998) found that orthographic overlap between the prime and target produces facilitation with a short SOA (e.g., 32 ms), but it gives way to inhibition as time passes (after 40 ms). On the other hand, phonological information begins to be activated a little later when the lexical competition at the level of orthography begins to produce inhibition. Thus, according to this account, the absence of facilitation for homophones observed in L1 speakers in the present study is because the facilitation from the pre-activation of the phonology of the target by the prime was cancelled out by the inhibition occurring from the orthographic representations. At the same time, the reason for stronger form facilitation for homophone and near-homophone pairs in L2 speakers could be because inhibition at the level of orthography is weak in the L2. Furthermore, the finding that more than 80% of the stimuli were indicated to be known and were correctly translated may suggest that the orthographic representations of at least some L2 word stimuli might not be fuzzy.

Although which mechanism indeed operates for the form facilitation for homophone or near-homophone pairs is unclear, the finding that prime-target pairs without a nonnative contrast (e.g., *dear-TEAR*) did not produce strong orthographic facilitation for the L2 participants in the current study may need more attention. This is because Jiang and Pae (2020) suggest that the levels of orthographic form confusion (for word pairs that are not phonologically related) could differ depending on L2 participants' native language. In Jiang and Pae's study (2020), the researchers let both L1 and L2 participants perform an LDT. In the stimuli they used, some of their nonwords were pseudo-homophones (e.g., *bote* which originated from *boat*), and some nonwords were orthographic control items, which were different from the pseudo-homophones by one letter and were not homophonic to real words (e.g., *wote* which originated from *bote*). Jiang and Pae (2020) found that L2 speakers whose first languages were alphabetic (e.g., Spanish or French) showed slower reaction times to homophonic nonwords (e.g., *bote*) just like L1 speakers. This suggests that these L2 speakers had difficulty when identifying the pseudo-homophones as nonwords because the phonological information of these items (e.g., /boʊt/), which sounded like real words, interfered with their decisions. However, another group of L2 speakers whose first languages were non-alphabetic (Chinese and Japanese), did not show this pseudo-homophone effect in reaction times. Instead, they showed higher error rates on pseudo-homophones than L1 speakers and the other group of L2 speakers with alphabetic first languages, indicating that non-alphabetic L2 speakers frequently identified pseudo-homophones as real words during the LDT. These researchers interpreted that this latter finding might have come from orthographic form confusion. Note that the pseudo-homophones (e.g., *bote*) look like more real words (e.g., *boat*) than orthographic controls (e.g., *wote*).

The Jiang and Pae (2020) study proposes that Korean L2 speakers, whose L1 is an alphabetic and orthographically shallow language, may depend more on phonology for lexical access whereas other L2 speakers whose first languages are non-alphabetic depend more on orthography. This is because L2 speakers often transfer their reading strategy in their L1 to reading in the L2 (Wang et al., 2003). Interestingly, most previous studies that report significant facilitative orthographic neighbor priming in the L2 (Jiang, 2021; Nakayama & Lupker, 2018; Qiao & Forster, 2017) have been conducted with Chinese or Japanese L2-learner-of-English participants. On the other hand, when Korean L2 learners of English were recruited for Park's study (2021), these participants did not show significant facilitation from the orthographic neighbor prime, even though he used stimuli that retained very similar linguistic properties to those used in Nakayama and Lupker's study (2018). Park (2021) initially interpreted that this finding was probably because the participants in his study had more precise orthographic representations than those in the Nakayama and Lupker study (2018). However, the proficiency of L2 participants in Park's study was actually comparable to those in the Nakayama and Lupker study (2018) when judged based on the participants' TOEIC scores. Furthermore, the current study replicated Park's finding that native Korean-speaking L2 speakers, even those with low proficiency, did not show facilitative orthographic neighbor priming. Therefore, it seems to indicate that form facilitation from orthographic form overlap in the L2 may not be very strong in native Korean speakers. If that is indeed the case, it can be assumed that facilitative orthographic neighbor priming may be language specific. In other words, whether form facilitation is driven by orthography or phonology may depend on what L2 speakers' first language is.

This is not to say that orthographic form priming does not occur at all for L2 speakers of alphabetic L1 backgrounds because significant orthographic form facilitation has been observed, even in this population when stimuli with other types of orthographic form overlap were used (e.g., *freeze-FREE* in Diependaele et al., 2011 or *example-exam* in Heyer & Clashen, 2015). The argument this study makes is that the magnitude of orthographic form facilitation may be weaker for L2 speakers of orthographically shallow, alphabetic L1s compared to that for other L2 speakers of logographic L1s.

5.3. The Role of Orthographic Precision in Form Priming

The present study found that orthographic precision did not modulate inhibitory form priming in the L1. This finding is contradictory to Andrews and Hersch's finding (2010) that native English-speaking participants' orthographic precision modulated the patterns of form priming in a masked priming LDT. Based on that finding, these researchers suspected that when participants had precise orthographic representations, they might be able to accumulate sufficient evidence to activate the prime even within a short SOA (50 ms). In consideration of their interpretation, it can be reasoned that the reason for a null effect of orthographic precision on form priming found in the L1 speakers in the current study might be due to the longer SOA (67 ms) employed by the present study. Within this long SOA, all L1 speakers who had more precise and larger lexicons than L2 speakers might be able to activate the prime. One may point out that these L1 speakers still did not show significant form inhibition even with this long SOA. This is suspected to be because of the use of short stimuli. As mentioned above, the mean length of the stimuli used in the present study was shorter than those in other studies. When nonwords are created by switching one letter from real words, and if stimuli are short, discrimination of word targets from nonword foils is easy. Previous research suggests that when the discrimination is easy, lexical competition is weak (De Moor et al., 2007; Forster & Veres, 1998).

The more important finding for the purposes of the present study is that orthographic precision modulated the strength of form facilitation in the L2. This finding is consistent with Andrews and Hersch's finding (2010). However, one question arises if we pay our attention to the finding that orthographic precision more strongly affected form facilitation under the /l/-/ɹ/ contrast condition (see Table 18 and Figure 7). How could orthographic precision modulate the priming that appears to have come from phonology?

The first conceivable scenario is to assume that the locus of the significant form facilitation detected under the /l/-/ɹ/ contrast condition was fuzzy orthographic representations rather than phonology. As discussed several times earlier, the current study found that (1) words with nonnative contrasts were more difficult for initial form-meaning mappings and (2) orthographic form-meaning mappings for these words were fuzzier than other words without a nonnative contrast. It is also worth noting that in Park's study (2021), in a mid-term priming experiment in which long words were used as stimuli (six- to ten-letter words), orthographic precision significantly modulated facilitative orthographic form priming. On the contrary, in a masked priming experiment in which short stimuli (four- to five-letter words without a difficult phonological contrast) were used, orthographic precision did not significantly modulate the strength of form facilitation. Because the acquisition of spellings of long words will be more challenging than short words, the findings of Park's study (2021) may suggest that the role of orthographic precision is more readily observable for words that are difficult to precisely encode.

Similarly, in the present study, the role of orthographic precision was observed for minimal pairs with the /l/-/ɹ/ contrast that should be more troublesome for orthographic encoding. At the same time, Figure 7 shows that orthographic precision affected form facilitation at least to a certain extent under the homophone condition, though this impact was not significant when statistically tested. On the other hand, orthographic encoding of the short words without nonnative contrasts that were used in the current study might not be very difficult. Therefore, this could be a reason why orthographic precision modulated form priming only under the /l/-/r/ contrast condition but not under the /f/-/p/ and the /b/-/t/ or /d/-/t/ conditions.

A second scenario can be proposed based on the assumption that form facilitation was produced by an overlap in phonemes between the prime and the target. In this case, as the orthographic representations became more precise, any facilitation coming from the pre-activation of sub-lexical components of the prime might have become less observable due to stronger lexical competition originating from orthographic representations. Under this scenario, the reason why a relatively stronger influence of orthographic representations was observed only under the homophone and the /l/-/ɹ/ contrast conditions was because form facilitation was strong only under these two conditions whereas it was weak under the other prime conditions. As reviewed in Section 2.2.2, unlike in perceptual identification or naming tasks, phonological priming for L1 speakers in visual LDTs has usually been fragile (Bowers et al., 1998; Davis et al., 1998; Kinoshita & Norris, 2012a; Rastle & Brysbaert, 2006; Shen & Forster, 1999) even though could occasionally be significant (Ferrand & Grainger, 1992; Grainger & Ferrand, 1994; Rastle & Brysbaert, 2006). Therefore, it is reasonable to suspect that lexical competition at the level orthography might have prevented any facilitative effects from the prime even in L2 speakers just as in L1 speakers.

Finally, it may be possible to guess that the L2 participants who used an alphabetic and orthographically shallow first language depended less on phonology as they developed stronger orthographic representations. Of course, whether proceeding from a written word to the lexical entry for that word uses a phonological conversion of the letters in visual word recognition has been a controversial topic for decades. Many researchers have claimed that lexical access to written words does not necessarily require phonological recoding (e.g., Coltheart, 2006; Coltheart et al., 1977; Davis et al., 1998; Martensen, Dijkstra, & Maris, 2005). Nevertheless, Davis and his colleagues (1998) report that in a masked phonological priming LDT, some fourth-grade native-English L1 child participants, who were slower in lexical decisions (but not those who were faster in lexical decisions), appeared to have responded faster when the target (e.g., *wash*) was primed by its pseudo-homophone (e.g., *wosh*) compared to when it was primed by an orthographic control prime (e.g., *wesh*). In addition, it should be noted that in Davis and his collaborators' study (1998), the participants were native English speakers whose L1 was an orthographically deep language. Some researchers hypothesized (Andrews, 1997; Brysbaert, forthcoming) that speakers of orthographically deep languages such as English would retain more precise orthographic representations which lead to direct access to the meaning from the orthography because depending on the phonological route for lexical access is inefficient for them. Thus, phonological recoding would be more likely to occur in L2 speakers if their native language has regular spelling-pronunciation mappings as in Korean because they are used to taking the phonological route.

Finally, before concluding this section, a surprising finding of the present study needs to be discussed. The finding that orthographic precision significantly modulated the form *inhibition* for L2 speakers under the same initial overlap condition (e.g., *play-PLAN*) (see Table 18 and

Figure 7) is particularly interesting with regard to the findings of a recent study (Jiang & Wu, 2022), in which Chinese L2 learners of English showed significant orthographic form facilitation for prime-target pairs with orthographic overlap at both word-initial (e.g., *rubber-RUB*) and word-final positions (e.g., *stage-AGE*). In the current study, the low-proficiency L2 speakers showed marginally significant inhibition if prime-target pairs shared the same initial (e.g., *play-PLAN*), whereas they showed a trend toward facilitation if prime-target pairs differed in the first letter (*dear-TEAR*) (Table 17). This suggests that the strong form facilitation observed regardless of the overlap position in Chinese L2 learners of English may not be a universal phenomenon.

In an attempt to determine the reason for the inhibitory trend observed in low-proficiency L2 speakers under the same initial overlap condition, additional exploratory analyses were performed. In these analyses, the magnitudes of form priming were compared after separating the prime-target pairs with medial and final non-overlap. As noted earlier, form inhibition has been observed in auditory priming studies if the prime and target shared the initial three phonemes, not only in the L1 (Dufour & Peereman, 2003b; Gor & Cook, 2020; Slowiaczek and Hamburger, 1992) but also in the L2 (Gor & Cook, 2020; Cook & Gor, 2015). Therefore, it was thought possible that the marginally significant inhibition for low-proficiency L2 speakers might have occurred due to their greater dependence on phonology and lexical competition triggered by an initial three-phoneme overlap. If so, the inhibition that these participants showed for prime-target pairs which differed in the last letter was expected to be stronger than that for other pairs that differed in a middle letter. However, during the analyses, it was found that many prime-target pairs that differed in the last letter did not share the same initial three phonemes either because whole words often consisted of only two (e.g., *bay* or *eat*) or three phonemes (e.g., *tape–TALE*) or because prime-target pairs sometimes did not share the same vowel (e.g., *play-PLAN*).

Therefore, in the second analysis, we compared the patterns of form priming between pairs sharing the body (the onset [initial consonant or consonant cluster] + vowel) (e.g., *feet-FEED*) and other pairs without a shared body (e.g., *then-TEEN*). Based on a previous study that reports Korean native speakers' greater dependence on the body than the rime (vowel + consonant[s]) in visual recognition of Korean words (Kim & Bolger, 2016), it was suspected to be plausible that low-proficiency L2 speakers might have under-processed the rime when the prime and target differed only in the last letter because they transferred their reading strategy for L1 words to L2 word reading. In this case, unresolved confusion of the target with the prime could lead to inhibition, which indicates transient confusion that takes time to resolve (Cook, Pandža, Lancaster, & Gor, 2016). Therefore, prime-target pairs sharing the body were predicted to show stronger form inhibition in low-proficiency L2 speakers than other pairs without a shared body.

However, contrary to the predictions, the results indicated that low-proficiency L2 speakers did not show stronger inhibition when the prime and target differed only in the last letter than when they differed in a middle letter, or when the prime-target pairs shared the body than when they did not. (See Appendix L for the specific results from the analyses). Nonetheless, it should be remembered that these analyses were performed completely for an exploratory purpose. For that reason, interpreting these results requires great caution. Because the linguistic properties (length, frequency and neighborhood density) of the stimuli used for the two conditions were not controlled (either in Analyses 1 or 2), fair comparisons between the two conditions (medial vs. last non-overlap or with vs. without a shared body) should be limited.

Interestingly, a similar finding that showed stronger inhibition when the prime and target shared the initial two or three letters was observed in previous visual word recognition experiments (Colombo, 1986). Colombo (1986) documented that in unmasked form priming experiments, native-Italian speakers showed significant inhibition when the Italian word prime and target overlapped the initial two or three letters, whereas they showed weak inhibition when the prime and target shared rhymes (i.e., last letters) unless they were high-frequency words. She thought that this was because under the former condition, the target probably began to interact with the prime earlier than under the latter condition. This mechanism is consistent with the so-called *early selection* in spoken word recognition (Marlen-Wilson, 1987) which refers to the identification of words before sufficient acoustic-phonetic information has become available (i.e., when only an initial part of words is heard). Therefore, the findings of Colombo's study (1986) appear to suggest that the dependence on phonology during visual word identification may indeed occur and be observable in a priming experiment if participants' first language is an orthographically shallow language such as Italian. Alternatively, the results of Colombo's study (1986) can be explained by the form confusion caused by under-processed rime as described in the previous paragraph. However, since Colombo (1986) did not closely inspect the number of shared phonemes between the prime and target, or whether the prime and target shared the body, the real reason for the inhibition for prime-target pairs sharing the same initial two or three letters that was observed in her study remains unknown.

Although we cannot be sure why the inhibition for prime-target pairs with an initial letter overlap seemed to be stronger for L2 speakers with relatively poorer orthographic representations, it is worth mentioning that in the current study, under the initial overlap condition, L1 speakers showed slightly weaker form inhibition compared to that under the other prime conditions (see also Frisson et al., 2014 for the same results obtained from native English-speaking L1 participants). Thus, weaker inhibition under this prime condition indicates native-like processing. In consequence, it is reasonable to conclude that this study has found that as L2 speakers develop orthographic representations, they display a nativelike pattern of priming (i.e., a null priming effect) even under this condition as well as under the other prime conditions. Future studies are necessary to clearly determine whether Korean L2 learners of English without sufficient L2 input indeed produce inhibition when the prime and target share the initial letter(s) and why this phenomenon occurs.

5.4. The Role of Vocabulary Size

In the current study, the influence of vocabulary size on form priming for all words was non-significant. However, when its impact on form priming for four-letter words was examined separately, the effect of vocabulary size was found significant. According to the lexical tuning hypothesis, (Castles et al, 1999; Castles et al., 2007; Forster & Veres, 1998; Qiao & Forster, 2013), as vocabulary size grows, words become better tuned in order not to be confused with their neighbors. Thus, vocabulary size is assumed to contribute to the development of precise orthographic representations. However, in the present study, orthographic precision was controlled for by spelling task scores. Therefore, the reason for its significant effect on the form priming that was observed in a model where spelling task scores were fitted as a covariate cannot be because vocabulary scores were related to orthographic precision. Given this situation, interaction-based models (Nakayama et al., 2008; Segui & Grainger, 1990) provide an alternative explanation. According to this view, it is assumed that when a prime with many neighbors is encountered, its neighbors are coactivated altogether, and they collectively produce large inhibition for the activation of the target. It is also thought that because it is more likely for the target to have a stronger competitor (i.e., its high-frequency neighbor) when its neighborhood size is big, the strong influence of the neighborhood size on form priming is understood to come from the activation of the strong competitor by the prime, which inhibits target identification.

Given this account based on the interaction-based models (Nakayama et al., 2008; Segui & Grainger, 1990), the reason why the effect of vocabulary size (i.e., the neighborhood density of stimuli) was non-significant for three-letter stimuli can be inferred to be because these shorter stimuli could not activate their neighbors. Based on the assertion that an initial three-phoneme overlap is required for activating the competitors (Dufour & Peereman, 2003b; Gor & Cook, 2020; Slowiaczek & Hamburger, 1992) in spoken word identification, it is arguable that at least a three-letter overlap would be required to detect a reliable influence of the neighbors, even in written word recognition. On the other hand, considering that the studies that reported a significant interaction between form priming and neighborhood density always used stimuli with a wide range of the orthographic neighborhood density (e.g., Andrews & Hersch, 2010; Davis & Lupker, 2006; Forster et al., 1987; Nakayama et al., 2008), the reason for the null effect of vocabulary size on form priming for five-letter stimuli is considered to be because they did not have a sufficient number of neighbors.

If the view that vocabulary size substantially affects form priming can be accepted, it can also explain why L1 speakers usually show stronger lexical competition than L2 speakers (e.g., Jiang, 2021; Nakayama & Lupker, 2018; Qiao & Forster, 2017). In the current study, although limited due to the small sample size, it was found that L2 speakers with large and precise mental lexicons can show a nativelike pattern of form priming. Therefore, the present study supports the idea that when controlling for vocabulary size, L1 and L2 lexicons are similar (Brysbaert et al., 2017).

Before wrapping up this section, it is necessary to note that the influence of vocabulary size may be more difficult to find in same-script bilinguals. This is because previous research suggests that when encountering an L2 word stimulus, even its L1 neighbors are activated (e.g., Bijeljac-Babic et al., 1997; Van Heuven et al., 1998). Therefore, future studies exploring the role of L2 vocabulary size would need to find a way to control the influence of cross-language orthographic neighbors.

5.5. The Relationship between Phonological Priming and Sound Perception Ability

Based on the difficulty of L2 speakers accurately identifying the /l/-/r/ contrast observed in the auditory tasks, it may be thought that the origin of the strong facilitation for prime-target pairs with this contrast would be the fuzzy phonological representations. However, this study found no evidence that the improvement in sound perception ability contributes to weaker form facilitation. Nonetheless, we do not think that the latter finding threatens our interpretation that fuzzy phonological representations of the /l/-/ɹ/ contrast is the fundamental source of strong form facilitation for words with this contrast. This is because the stimuli used under the /l/-/ɹ/ contrast prime condition had linguistic properties similar to those used under the other prime conditions. Therefore, if it was not the fuzzy phonological representations, there is no other theoretical reason for this contrast to produce stronger facilitation. However, it should be explained why the development of sound perception ability did not account for the weaker form facilitation.

The most plausible reason is that the sound identification tasks used in the current study were not very suitable to evaluate the quality of phonological representations. In other words, even though these tasks were able to show that L2 speakers were less accurate in sound perception of confusing contrasts compared to L1 speakers, they might not be able to discriminate L2 speakers with more precise phonological representations from those with poorer representations. This is because participants could have sufficient time to make use of their explicit knowledge of L2 phonemes and to carefully assess the properties of the aural stimuli based on their knowledge. Supporting this possibility, the reaction time data from the word and phoneme identification tasks (which are not reported in Chapter 4) showed that L2 speakers' responses to stimuli with the nonnative contrasts were much slower (110 ms for the /f/-/p/ and 227 ms for the /l/-/ɹ/ contrast) than those with the other contrasts (the /b/-/t/ and /d/-/t/ contrasts). L1 speakers did not show such a big gap as L2 speakers. (They were 37 ms slower for the /f/-/p/ and 32 ms for the /l/-/ɹ/ contrasts compared to their reaction times to the /b/-/t/ and /d/-/t/ contrasts.) These findings suggest that participants' explicit knowledge of L2 phonemes, as well as phonological representations, might have played a considerable role during L2 speakers' sound identification.

5.6. Limitations and Future Studies

First, as noted above, this study did not observe a meaningful relationship between sound perception ability and lexical competition. However, we consider that this might be because of methodological limitations. Therefore, it is still undetermined whether accurate sound perception indeed does not affect lexical competition. Although we could think of several possibilities for accurate measurement of the quality of phonological representations (such as the use of reaction time data from auditory phonological priming experiments or the visual word paradigm), we are not sure whether the amount of auditory priming or eye-tracking data could be used as a reliable independent variable. Future research exploring the interaction between sound perception and lexical competition should first develop a good measure of phonological representations.

Second, this study utilized a between-item design. Thus, the target was primed by a word that was used only under a particular prime condition but not by other words from other prime conditions. Therefore, we admit that although we controlled the length, frequency and neighborhood density of stimuli, the findings could have been affected by other uncontrolled characteristics of stimuli such as such as semantic concreteness, cognateness (e.g., De Groot & Keijzer, 2000) or contextual or semantic diversity (e.g., Hamrick & Pandža, 2020). In addition, in the post-hoc analyses (Section 4.7 and 4.8), even the three linguistic properties (length, frequency and neighborhood density) were not controlled. Therefore, making fair inferences from these post-hoc analyses should be limited.

In future studies, the influence of the L1 in orthographic and phonological priming in visual word recognition needs to be explored. The current study found some evidence suggesting that the strong orthographic priming observed in previous studies with native Chinese- and Japanese-speaking L2 learners of English (Jiang, 2021; Nakayama & Lupker, 2018; Qiao & Forster, 2017) may not be uniformly found in other bilinguals if their L1 is not a non-alphabetic language. At the same time, this study raises a question about whether comparable phonological priming would be observable from other L2 learners whose first language is non-alphabetic (although cross-language homophone priming studies (see Section 2.2.4) suggest that a certain level of facilitation from the phonological prime may also be observable even in this population). It is also possible that the strength of orthographic or phonological priming could differ depending on not only the depth of orthography of first languages but also that of second languages (e.g., English vs. Italian) (Andrews, 1997; Brysbaert, forthcoming). These questions may be important for estimating the generalizability of any findings of second language research.

5.7. Pedagogical Implications

The primary purpose of the current study was to explore the nature of L2 form representations rather than to draw pedagogical implications. However, the findings of this study provide some suggestions for language instruction. The present study proposes that when the phonological forms of L2 words involve a difficult phonological contrast, the acquisition of accurate spellings, as well as pronunciation of these words may be more challenging than that of the L2 words without a confusing contrast. Although it cannot yet be concluded that the stronger form facilitation that was observed for the prime-target pairs with a difficult contrast in the LDT was driven by orthographic form confusion, the analyses of L2 speakers' translation responses and familiarity ratings indicate that the initial stage of orthographic form-meaning mappings of L2 words that are differentiated by a difficult contrast is likely to be more problematic. Therefore, L2 learners would need to pay additional attention to the spelling if the L2 words have a confusing phonological contrast.

Chapter 6. Conclusion

The current study was conducted to track the developmental trajectory of L2 orthographic representations. A crucial role of the quality of orthographic representations in determining the patterns and the strength of orthographic neighbor priming was detected. In addition, as predicted, the challenge that a confusing phonological contrast poses for accurate L2 word form encoding was observed. Vocabulary size was also judged to play some role in lexical competition, and a significant impact of this construct on orthographic form facilitation for four-letter words was observed. On the other hand, global proficiency did not greatly influence orthographic neighbor priming, indicating that global proficiency and orthographic precision are distinct constructs, and that the latter is a better predictor for lexical competition. All these findings support the FLR hypothesis which argues for a developmental nature of lexical representations and understands lexical competition as the product of sufficiently specified orthographic representations.

However, several important questions remain unanswered. The first question is the exact locus of the stronger form facilitation for homophones or near-homophones. Even though important previous research on masked phonological priming and the automatic activation of phonological codes during visual word recognition can lead us to a phonology-based interpretation, the present study's claim that the fuzzier orthographic representations of homophones or near-homophones is a source of stronger form facilitation for these words should not be disregarded. Another unresolved question is whether the development of phonological representations of words with a confusing contrast can lead to stronger lexical competition between these words during written word identification. Developing a good measure of the precision in phonological representations may help answer both questions.

Appendix A: Lexical Decision Task Items

Homophone			Words with the /l/-/r/ contrast			Words with the /f/-/p/ contrast		
Related prime	Unrelated prime	Target	Related prime	Unrelated prime	Target	Related prime	Unrelated prime	Target
too	say	TWO	right	where	LIGHT	full	sick	PULL
son	job	SUN	ray	box	LAY	fool	card	POOL
see	did	SEA	road	mine	LOAD	four	each	POUR
meet	kids	MEAT	red	top	LED	fit	leg	PIT
steal	price	STEEL	lap	net	RAP	fan	nor	PAN
heal	boot	HEEL	royal	seats	LOYAL	pork	tale	FORK
toe	sin	TOW	lack	bend	RACK	fond	eats	POND
pole	hers	POLL	rid	van	LID	feel	guys	PEEL
one	yes	WON	read	wish	LEAD	fast	ride	PAST
here	just	HEAR	rock	test	LOCK	fair	song	PAIR
sell	ring	CELL	late	work	RATE	found	ready	POUND
week	hurt	WEAK	river	cross	LIVER	file	form	PILE
sight	waves	SITE	rip	bow	LIP	face	lost	PACE
pray	whom	PREY	law	key	RAW	paint	tells	FAINT
flee	trim	FLEA	laid	sees	RAID	fine	must	PINE
seem	luck	SEAM	race	spot	LACE	pin	sum	FIN

Words with the /b/-/t/ or /d/-/t/ contrast			Word pairs sharing the initial		
Related prime	Unrelated prime	Target	Related prime	Unrelated prime	Target
down	make	TOWN	then	call	TEEN
teach	likes	BEACH	tape	wind	TALE
try	lot	DRY	feet	boss	FEED
ball	busy	TALL	lake	post	LANE
bone	lily	TONE	eat	sit	EAR
bag	cop	TAG	big	has	BUG
team	king	BEAM	sing	gold	SINK
tip	lab	DIP	model	finds	MOTEL
took	used	BOOK	been	look	BEER
die	pay	TIE	home	away	HOLE
tell	from	BELL	black	daddy	BLOCK
dear	boys	TEAR	may	old	MAP
boy	bad	TOY	share	dying	SHAPE
dug	per	TUG	bay	pen	BAT
train	state	DRAIN	hate	pick	HATS
tent	lick	DENT	fly	gas	FLU

Word prime-nonword targets					
Related prime	Unrelated prime	Target	Related prime	Unrelated prime	Target
let	his	DET	her	who	HIR
them	back	SHEM	think	would	GHINK
our	way	OUS	said	over	TAID
guy	new	KUY	stop	nice	STOD
name	told	NAWE	kind	girl	KUND
ask	dad	OSK	men	own	MES
saw	mom	GAW	whole	bring	WHOLA
idea	many	IDEI	hard	most	RARD
yet	use	YIT	hope	says	FOPE
few	kid	GEW	turn	gone	TURB
case	hand	CESE	such	year	SICH
hit	ago	VIT	days	true	ZAYS
game	word	VAME	eyes	shot	EYEF
goes	half	MOES	walk	fire	GALK
gun	set	YUN	cool	rest	COOB
city	glad	TITY	hour	safe	HOIR
air	win	AOR	sent	lord	SEST
beat	drop	BEAH	act	lie	UCT
fall	none	SALL	wear	kiss	WEAF
wake	bill	WAJE	ship	blue	GHIP
pain	club	HAIN	quit	blow	NUIT
stick	eight	SRICK	honor	relax	HOSOR
kept	bank	LEPT	rich	dark	XICH
cute	land	CUSE	bar	fix	BOR
army	cops	ARHY	fat	ran	FUT
born	fish	BORB	harry	smell	HAFRY
enjoy	store	ENDOY	list	loud	TIST
ice	art	IKE	bus	age	BIS
liked	group	LOKED	brain	chief	SRAIN
peace	favor	BEACE	desk	cook	DESR
oil	due	OIT	heat	join	CEAT
wet	pig	WEM	fake	unit	ZAKE
meal	dump	YEAL	lover	rooms	LOHER
cow	egg	XOW	odd	pet	ODS
ill	spy	ULL	minds	bored	MUNDS
cats	role	CAKS	kit	cap	XIT
term	flag	RERM	stake	humor	SBAKE
aid	zoo	AIV	jam	hut	NAM
lamp	leak	MAMP	task	dive	QASK
link	fits	LONK	core	oath	CIRE

Appendix B: Spelling Task Items (Adopted from Andrews et al., 2020)

1. Spelling Dictation Task

A native speaker reads the word aloud to the participant followed by the sentence containing the word. The participant then types in the spelling of each target word.

1. ABSTINENCE
The ex-alcoholic found it very difficult to maintain complete abstinence from drinking.

2. ACQUAINTANCE
She knew the woman as an acquaintance, but she was not a close friend.

3. DIGESTIBLE
The nurses had to blend the food to make it digestible for the patient.

4. CONCILIATORY
She tried to adopt a conciliatory approach to avoid further conflict.

5. PISTACHIO
The biscuit with pistachio nuts was delicious.

6. WARRANTY
The TV seemed a good buy because it had a three-year warranty.

7. RHEUMATIC
The old woman suffered rheumatic pain in all her joints.

8. CRESCENDO
The music reached a crescendo towards the end of the symphony.

9. ASYMMETRY
She found the asymmetry of the design very appealing.

10. AFFLUENT
By comparison with most Asian countries, Australia is very affluent.

11. DILIGENT
Most of the students are lazy but this boy is very diligent in completing his work.

12. AGGRAVATION
The noise from the next classroom was a constant aggravation to the teacher.

13. COLLOQUIAL

It is usually not appropriate to use colloquial language in an essay.

14. EUPHORIC
The student felt euphoric when she completed her last exam.

15. BROCCOLI
Children often dislike broccoli and other green vegetables.

16. SOMERSAULT
The gymnast turned a somersault before landing back on the bar.

17. OBLIVION
Many sportsmen achieve fame when they are young but then sink into oblivion.

18. RHYTHMICAL
The rhythmical beat of the drums was mesmerizing.

19. RAVENOUS
She had missed lunch so by dinner time she felt ravenous.

20. PERSUADE
She tried to persuade him to her point of view.

2. Spelling Recognition Task

The participant is given unlimited time to decide whether each word is spelled correctly or wrongly.

attitude, criticism, benafit, refrences, misary, psycology, political, glamourous, reciept, available, addmission, tounge, appreciate, materilistic, independent, chronicle, seperate, senior, behavior, atterney, sufficient, efficiency, implie, courtesy, mortgage, govenment, basicly, privalege, consequence, sieze, suspicious, prosedure, conveinient, insurance, imminant, guitar, elementary, sacrifice, commitment, decrepit, jeapordize, forfeit, fulcrum, annihlate, distinguish, inquirey, sincirely, equivical, gaurantee, delecate, bachelor, annual, necesscarily, favorate, announcment, severe, occurence, insatiable, partitionining, asure, exhibition, warrent, interrogate, havoc, conscientious, parallel, interpretation, bureaucracy, importent, negotiate, proliferate, vigilent, missellaneous, curiculum, plagarism, acomplice, pollution, permanent, aplause, subpoena, accommodation, attentsion, rendezvous, subtlety, honerable, inhibition, classafied, assessor

Appendix C: Vocabulary Test Items (Adopted from Shipley, 1940)

Word	Options			
(1) TALK	draw	eat	<u>speak</u>	sleep
(2) PERMIT	<u>allow</u>	sew	cut	drive
(3) PARDON	<u>forgive</u>	pound	divide	tell
(4) COUCH	pin	eraser	<u>sofa</u>	glass
(5) REMEMBER	swim	<u>recall</u>	number	defy
(6) TUMBLE	drink	dress	<u>fall</u>	think
(7) HIDEOUS	silvery	tilted	young	<u>dreadful</u>
(8) CORDIAL	swift	muddy	leafy	<u>hearty</u>
(9) EVIDENT	green	<u>obvious</u>	skeptical	afraid
(10) IMPOSTER	conductor	officer	book	<u>pretender</u>
(11) MERIT	<u>deserve</u>	distrust	fight	separate
(12) FASCINATE	welcome	fix	stir	<u>enchant</u>
(13) INDICATE	defy	excite	<u>signify</u>	bicker
(14) IGNORANT	red	sharp	<u>uninformed</u>	precise
(15) FORTIFY	submerge	<u>strengthen</u>	vent	deaden
(16) RENOWN	length	head	<u>fame</u>	loyalty
(17) NARRATE	yield	buy	associate	<u>tell</u>
(18) MASSIVE	bright	<u>large</u>	speedy	low
(19) HILARITY	<u>laughter</u>	speed	grace	malice
(20) SMIRCHED	stolen	pointed	remade	<u>soiled</u>

Word	Options			
(21) SQUANDER	tease	belittle	cut	<u>waste</u>
(22) CAPTION	drum	ballast	<u>heading</u>	ape
(23) FACILITATE	<u>help</u>	turn	strip	bewilder
(24) JOCOSE	<u>humorous</u>	paltry	fervid	plain
(25) APPRISE	reduce	strew	<u>inform</u>	delight
(26) RUE	eat	<u>lament</u>	dominate	cure
(27) DENIZEN	senator	<u>inhabitant</u>	fish	atom
(28) DIVEST	<u>dispossess</u>	intrude	rally	pledge
(29) AMULET	<u>charm</u>	orphan	dingo	pond
(30) INEXORABLE	untidy	<u>involatile</u>	rigid	sparse
(31) SERRATED	dried	<u>notched</u>	armed	blunt
(32) LISSOM	moldy	<u>loose</u>	supple	convex
(33) MOLLIFY	<u>mitigate</u>	direct	pertain	abuse
(34) PLAGIARIZE	<u>appropriate</u>	intend	revoke	maintain
(35) ORIFICE	brush	<u>hole</u>	building	lute
(36) QUERULOUS	maniacal	curious	devout	<u>complaining</u>
(37) PARIAH	<u>outcast</u>	priest	lentil	locker
(38) ABET	waken	ensue	<u>incite</u>	placate
(39) TEMERITY	<u>rashness</u>	timidity	desire	kindness
(40) PRISTINE	vain	sound	<u>first</u>	level

Appendix D: Word and Phoneme Identification Task Items

Word identification task items				Phoneme identification task items			
light	right	lead	read	/ɹaɪb/	/laɪl/	/fæsp/	/pæsp/
pull	full	fast	past	/ɹeɪb/	/leɪb/	/fɛv/	/pɛv/
down	town	took	book	/ɹoʊk/	/loʊk/	/faʊmd/	/paʊmd/
ray	lay	rock	lock	/ɹɛŋ/	/lɛv/	/faɪʃ/	/paɪf/
fool	pool	pair	fair	/læŋ/	/ɹæf/	/feɪp/	/peɪp/
teach	beach	die	tie	/ɹɔɪəf/	/lɔɪəf/	/peɪnk/	/feɪnk/
load	road	late	rate	/læl/	/ɹæl/	/faɪg/	/paɪg/
pour	four	found	pound	/ɹɪʃ/	/lɪθ/	/pɪm/	/fɪm/
try	dry	tell	bell	/ɹit/	/lim/	/daʊm/	/taʊm/
red	led	river	liver	/ɹɑʃ/	/lɑʃ/	/tif/	/biθ/
fit	pit	pile	file	/leɪf/	/ɹeɪf/	/tɹaɪv/	/dɹaɪf/
ball	tall	dear	tear	/ɹɪvəd/	/lɪvəm/	/bɔv/	/tɔv/
lap	rap	rip	lip	/ɹɪn/	/lɪf/	/boʊm/	/toʊθ/
fan	pan	face	pace	/lɔtʃ/	/ɹɔp/	/bæf/	/tæf/
bone	tone	boy	toy	/leɪp/	/ɹeɪm/	/tiʃ/	/biʃ/
royal	loyal	law	raw	/ɹeɪtʃ/	/leɪtʃ/	/tɪd/	/dɪθ/
pork	fork	paint	faint	/fʊp/	/pʊŋ/	/tʊf/	/bʊv/
bag	tag	dug	tug	/fʊz/	/pʊz/	/daɪp/	/taɪn/
lack	rack	laid	raid	/fɔʃ/	/pɔf/	/tɛv/	/bɛʃ/
fond	pond	fine	pine	/fɪtʃ/	/pɪv/	/dɪz/	/tɪb/
team	beam	train	drain	/fæp/	/pæb/	/bɔɪb/	/tɔɪb/
rid	lid	race	lace	/pɔrf/	/fɔrf/	/dʌʃ/	/tʌl/
peel	feel	pin	fin	/fɑnp/	/pɑnp/	/tɹeɪm/	/dɹeɪm/
tip	dip	tent	dent	/fiθ/	/pif/	/tɛnf/	/dɛnf/

Appendix E: Cloze Test (Adopted from Brown, 1980)

DIRECTIONS

1. Read the passage quickly to get the general meaning.
2. Write only one word in each blank next to the item number. Contractions are considered as one word.
3. Check your answers.

EXAMPLE: The boy walked up the street. He stepped on a piece of ice.

He fell (1) ***down*** but he didn't hurt himself.

MAN AND HIS PROGRESS

Man is the only living creature that can make and use tools. He is the most teachable of living beings, earning the name of Homo sapiens. (1)____ ever restless brain has used the (2)___ and the wisdom of his ancestors (3)____ improve his way of life. Since (4)_____ is able to walk and run (5)____his feet, his hands have always (6)____ free to carry and to use (7)_____ . Man's hands have served him well (8)_____ his life on earth. His development, (9)_____ can be divided into three major (10)____, is marked by several different ways (11)_____ life.

Up to 10,000 years ago, (12)____ human beings lived by hunting and (13)_____. They also picked berries and fruits, (14)_____ dug for various edible roots. Most (15)_____, the men were the hunters, and (16)_____ women acted as food gatherers. Since (17)_____ women were busy with the children, (18)_____ men handled the tools. In a (19)_____ hand, a dead branch became a (20)____ to knock down fruit or (21)____ for tasty roots. Sometimes, an animal

(22) ____ served as a club, and a (23) ____ piece of stone, fitting comfortably into (24) ____ hand, could be used to break (25) ____ or to throw at an animal. (26) ____ stone was chipped against another until (27) ____ had a sharp edge. The primitive (28) ____ who first thought of putting a (29) stone at the end of a (30) ____ made a brilliant discovery: he (31) ____ joined two things to make a (32) ____ useful tool, the spear. Flint, found (33) ____ many rocks, became a common cutting (34) in the Paleolithic period of man's (35) ____. Since no wood or bone tools

(36) ____ survived, we know of this man (37) ____ his stone implements, with which he

(38) ____ kill animals, cut up the meat, (39) ____ scrape the skins, as well as (40) ____ pictures on the walls of the (41) ____ where he lived during the winter.

(42) ____ the warmer seasons, man wandered on (43) ____ steppes of Europe without a fixed (44) ____, always foraging for food. Perhaps the (45) ____ carried nuts and berries in shells (46) skins or even in light, woven (47) ____ . Wherever they camped, the primitive people (48) ____ fires by striking flint for sparks (49) ____ using dried seeds, moss, and rotten (50) ____ for tinder. With fires that he kindled himself, man could keep wild animals away and could cook those that he killed, as well as provide warmth and light for himself.

Appendix F: Lexical Characteristics of Stimuli

		Homophone (*n* = 16)		/l/ - /r/ (*n* = 16)		/p/ - /f/ (*n* = 16)		/d/ - /t/ or /b/ - /t/ (*n* = 16)		Orthographic neighbors at do not differ in the first letter (*n* = 16)		Statistical test
Target	Length	3.75 (SD: 0.58)		3.81 (SD: 0.75)		3.94 (SD: 0.57)		3.75 (SD: 0.68)		3.83 (SD: 0.65)		ANOVA $F(4, 75) = .24$, $p = .91$
	Log F	3.13 (SD: 0.92)		2.98 (SD: 0.46)		2.89 (SD: 0.47)		3.06 (SD: 0.33)		3.02 (SD: 0.57)		Welch $F(4, 36.77) = .45$, $p = .91$
	ORTH N	12.44 (SD: 6.77)		16.31 (SD: 6.78)		14.06 (SD: 6.46)		14.81 (SD: 5.71)		15.63 (SD: 8.14)		ANOVA $F(4, 75) = .77$, $p = .55$
	Mean bigram F	3,381 (SD: 1,337)		3,768 (SD: 1,740)		3,836 (SD: 1,837)		3,362 (SD: 2,001)		3,960 (SD: 1,686)		Welch $F(4, 37.29) = .42$, $p = .80$
		Related (*week-WEAK*)	Unrelated (*hurt-WEAK*)	Related (*read-LEAD*)	Unrelated (*wish-LEAD*)	Related (*full-PULL*)	Unrelated (*sick-PULL*)	Related (*dear-TEAR)*	Unrelated (*boys-TEAR)*	Related (*play-PLAN*)	Unrelated (*best-PLAN*)	
Prime	Length	3.81 (SD: 0.66)	3.81 (SD: 0.66)	3.81 (SD: 0.75)	3.81 (SD: 0.75)	3.94 (SD: 0.57)	3.94 (SD: 0.57)	3.75 (SD: 0.68)	3.75 (SD: 0.68)	3.88 (SD: 0.72)	3.88 (SD: 0.72)	ANOVA $F(9, 150) = .16$, $p = 1.00$
	Log F	3.85 (SD: 0.95)	3.81 (SD: 0.95)	3.61 (SD: 0.60)	3.63 (SD: 0.59)	3.66 (SD: 0.58)	3.66 (SD: 0.59)	3.87 (SD: 0.62)	3.87 (SD: 0.64)	3.95 (SD: 0.59)	3.94 (SD: 0.61)	Welch $F(9, 61.05) = .72$, $p = .69$
	ORTH N	13.94 (SD: 7.31)	13.75 (SD: 6.95)	15.94 (SD: 6.82)	14.38 (SD: 6.88)	13.13 (SD: 5.24)	13.19 (SD: 5.48)	14.94 (SD: 6.27)	14.63 (SD: 8.15)	15.31 (SD: 6.78)	13.81 (SD: 7.07)	Welch $F(9, 61.05) = .32$, $p = .97$
	Mean bigram F	3,693 (SD: 2,008)	3,778 (SD: 2,862)	3,840 (SD: 1,771)	3,896 (SD: 2,460)	3,622 (SD: 1,951)	3,607 (SD: 1,377)	3,681 (SD: 2,011)	3,794 (SD: 2,346)	3,541 (SD: 2,036)	3,227 (SD: 1,770)	Welch $F(9, 60.99) = .15$, $p = 1.00$

Log F: Log SUBTLEX-US Frequency, ORTH N: Number of orthographic neighbors (neighborhood density), Mean Bigram F: Mean bigram frequency

Appendix G: Participants' Biographical Information

1. Age

Low-prof. L2 speakers	Mid-prof. L2 speakers	High-prof. L2 speakers	L1 speakers	Statistical test
24.07 (SD: 3.97)	25.47 (SD: 4.22)	35.03 (SD: 6.56)	23.93 (SD: 6.74)	Welch $F(3, 62.99) = 21.95$, $p < .001$
Pairwise comparisons	(1) L1 speakers vs. Low-prof. L2 speakers: -0.14 (p = 1.00) (2) L1 speakers vs. Mid-prof. L2 speakers: -1.54 (p = .72) (3) L1 speakers vs. High-prof. L2 speakers: -11.10 (p < .001) (4) Low- vs. mid-prof. L2 speakers: -1.40 (p = .55) (5) Low- vs. high-prof. L2 speakers: -10.96 (p < .001) (6) Mid- vs. high-prof. L2 speakers: -9.56 (p < .001) * Games-Howell tests were conducted for statistical analyses.			

2. Length of Education

Low-prof. L2 speakers	Mid-prof. L2 speakers	High-prof. L2 speakers	L1 speakers	Statistical test
15.23 (SD: 1.45)	15.47 (SD: 1.57)	19.47 (SD: 2.85)	16.0 (SD: 2.23)	Welch $F(3, 62.86) = 18.12$, $p < .001$
Pairwise comparisons	(1) L1 speakers vs. Low-prof. L2 speakers: 0.77 (p = .40) (2) L1 speakers vs. Mid-prof. L2 speakers: 0.53 (p = .71) (3) L1 speakers vs. High-prof. L2 speakers: -3.47 (p < .001) (4) Low- vs. mid-prof. L2 speakers: -0.24 (p = .93) (5) Low- vs. high-prof. L2 speakers: -4.24 (p < .001) (6) Mid- vs. high-prof. L2 speakers: -4.00 (p < .001) * Games-Howell tests were conducted for statistical analyses.			

3. Age of Acquisition

Low-prof. L2 speakers	Mid-prof. L2 speakers	High-prof. L2 speakers	Statistical test
8.00 (SD: 1.49)	7.77 (SD: 1.81)	9.17 (SD: 2.47)	Welch $F(2, 55.94) = 3.33$, $p = .04$
Pairwise comparisons	(1) Low- vs. mid-prof. L2 speakers: -0.24 (p = .85) (2) Low- vs. high-prof. L2 speakers: -4.24 (p = .08) (3) Mid- vs. high-prof. L2 speakers: -4.00 (p = .04) * Games-Howell tests were conducted for statistical analyses.		

4. Proficiency

Low-prof. L2 speakers	Mid-prof. L2 speakers	High-prof. L2 speakers	Statistical test
23.33 (SD: 5.09)	31.43 (SD: 7.31)	36.93 (SD: 8.55)	ANOVA $F(2, 87) = 27.64$, $p < .001$
Pairwise Comparisons	(1) Low- vs. mid-prof. L2 speakers: -8.10 (p < .001) (2) Low- vs. high-prof. L2 speakers: -13.60 (p < .001) (3) Mid- vs. high-prof. L2 speakers: -5.00 (p = .011) * Bonferroni-adjusted significance tests were conducted for statistical analyses.		

Appendix H: Output from Mixed-Effects Models for the Lexical Decision Task

1. Accuracy for Word Targets

Fixed effects				
Predictor	Estimate	*SE*	*z*	*p*
(1) The reference category for Group: L1 speakers				
(Intercept)	3.35	0.38	8.77	< .001***
Related	-0.29	0.43	-0.67	.503
Condition	0.05	0.11	0.47	.636
GroupL2	-0.88	0.28	-3.16	.002
Related:Condition	-0.04	0.13	-0.31	.755
Related:GroupL2	0.61	0.47	1.29	.198
Condition:GroupL2	0.11	0.07	1.50	.133
Related:Condition:GroupL2	0.01	0.15	0.05	.964
(2) The reference category for Group: L2 speakers				
(Intercept)	2.47	0.31	7.90	< .001***
Related	0.32	0.20	1.60	.110
Condition	0.16	0.09	1.75	.080
Related:Condition	-0.03	0.06	-0.54	.590
Random effects	Variance	SD		
Participant	0.45	0.67		
Item	1.14	1.07		

Notes. Formula (glmer): Acc ~ Related * Condition * Group + (1 | Participant) + (1 | Item),
Significance codes: '***' < .001, '**' < .01, '*' < .05

2. Response Latency to Nonword Targets

Fixed effects				
Predictor	Estimate	*SE*	*t*	*p*
(1) The reference category for Group: L1 speakers				
(Intercept)	-1502.17	41.13	-36.52	< .001***
Related	8.70	10.74	0.81	.418
GroupL2	245.49	46.21	5.31	< .001***
Related:GroupL2	-12.51	11.20	-1.12	.264
(2) The reference category for Group: L2 speakers				
(Intercept)	-1256.68	25.00	-50.28	< .001***
Related	-3.81	7.32	-0.52	.603
Random effects	Variance	SD	Corr	
Participant	47335	217.57		
Group	7252	85.16	-0.41	
Item	1739	41.71		

Notes. Formula (lmer): RT_Inverse ~ Related * Group + (1 | Participant) + (1 + Related | Item),
Significance codes: '***' < .001, '**' < .01, '*' < .05

3. Accuracy for Nonword Targets

Fixed effects				
Predictor	Estimate	*SE*	*z*	*p*
(1) The reference category for Group: L1 speakers				
(Intercept)	3.67	0.22	16.67	< .001
Related	-0.03	0.20	-0.17	.868
GroupL2	-0.39	0.21	-1.84	.065
Related:GroupL2	-0.13	0.23	-0.58	.563
(2) The reference category for Group: L2 speakers				
(Intercept)	3.28	0.16	20.96	< .001
Related	-0.16	0.10	-1.64	.101
Random effects	Variance	SD		
Participant	0.69	0.83		
Item	0.94	0.97		

Notes. Formula (glmer): Acc ~ Related * Group + (1 | Participant) + (1 | Item),
Significance codes: '***' < .001, '**' < .01, '*' < .05

Appendix I: Results of the Lexical Decision Task after Excluding Two Prime-Target Pairs with Heteronyms

1. L1 and L2 Speakers' Mean Reaction Times (in Milliseconds) and Error Rates (in Parentheses) under Different Prime Conditions in the Masked Priming Lexical Decision Task

	Homophone		/l/ - /ɹ/		/p/ - /f/		/d/ - /t/ or /b/ - /t/		Orthographic neighbors sharing the word initial		Nonwords	
	Related (*meet-MEAT*)	Unrelated (*kids-MEAT*)	Related (*read-LEAD*)	Unrelated (*wish-LEAD*)	Related (*full-PULL*)	Unrelated (*sick-PULL*)	Related (*dear-TEAR)*	Unrelated (*boys-TEAR)*	Related (*play-PLAN*)	Unrelated (*best-PLAN*)	Related (*name-MAME*)	Unrelated (*long-MAME*)
L1 speakers (n = 30)	657.0 (5.1%)	640.2 (4.7%)	687.5 (10.4%)	666.6 (5.9%)	671.7 (6.8%)	655.2 (5.2%)	660.4 (5.0%)	649.0 (4.1%)	655.8 (6.4%)	655.8 (4.2%)	700.8 (4.6%)	704.7 (4.4%)
	-16.8 (-0.4%)		-20.9 (-4.5%)†		-16.5 (-1.6%)		-11.4 (-0.9%)		0 (-2.2%)		3.9 (-0.2%)	
L2 speakers (n = 90)	780.5 (8.6%)	801.2 (12.3%)	795.0 (11.5%)	814.6 (11.9%)	791.8 (6.8%)	805.6 (7.1%)	767.8 (5.1%)	774.7 (8.7%)	807.5 (6.3%)	790.2 (7.2%)	856.2 (7..02%)	862.6 (6.2%)
	20.7* (3.7%)		**19.6**** (0.4%)		13.8 (0.3%)		6.9 **(3.6%)***		-17.3 (0.9%)		6.4 (-1.0%)	

Notes. Significance codes: '**' < .01, '*' < .05, †: p < .10

2. Three L2 Speaker Groups' Mean Reaction Times (in Milliseconds) and Error Rates (in Parentheses) under Different Prime Conditions in the Masked Priming Lexical Decision Task

	Homophone		/l/ - /ɹ/		/p/ - /f/		/d/ - /t/ or /b/ - /t/		Orthographic neighbors sharing the initial		Nonwords	
	Related (*meet-MEAT*)	Unrelated (*kids-MEAT*)	Related (*read-LEAD*)	Unrelated (*wish-LEAD*)	Related (*full-PULL*)	Unrelated (*sick-PULL*)	Related (*dear-TEAR*)	Unrelated (*boys-TEAR*)	Related (*play-PLAN*)	Unrelated (*best-PLAN*)	Related (*name-MAME*)	Unrelated (*long-MAME*)
Low-proficiency L2 speakers (*n* = 30)	804.7 (10.7%)	837.3 (12.4%)	791.8 (13.4%)	820.6 (13.6%)	807.2 (7.1%)	828.7 (8.5%)	779.8 (3.4%)	805.4 (10.6%)	841.8 (7.8%)	800.5 (8.5%)	899.5 (10.5%)	903.8 (8.0%)
	32.6* (1.7%)		**28.8*** (0.2%)		21.5† (1.4%)		25.6 **(7.2%)***		-41.3† (0.7%)		4.3 **(-2.5%)***	
Medium-proficiency L2 speakers (*n* = 30)	718.5 (9.8%)	745.9 (15.1%)	747.1 (12.8%)	773.9 (14.5%)	734.9 (6.1%)	748.4 (7.8%)	740.9 (9.2%)	724.2 (11.7%)	750.5 (6.4%)	734.0 (9.3%)	797.4 (7.2%)	797.3 (7.0%)
	27.4* (5.3%)†		**26.8*** (1.7%)		13.5 (1.7%)		-16.7 (2.5%)		-16.5 (2.9%)		-0.1 (-0.2%)	
High-proficiency L2 speakers (*n* = 30)	814.4 (2.5%)	819.6 (5.6%)	842.9 (8.4%)	847.7 (7.8%)	834.3 (5.7%)	838.6 (3.0%)	781.4 (2.7%)	792.4 (4.4%)	830.6 (4.7%)	828.7 (3.8%)	873.4 (3.8%)	886.2 (3.6%)
	5.2 (3.1%)†		4.8 (-0.6%)		4.3 (-2.7%)		11.0 (1.7%)		-1.9 (-0.9%)		12.8 (-0.2 %)	
L1 speakers (*n* = 30)	657.0 (5.1%)	640.2 (4.7%)	687.5 (10.4%)	666.6 (5.9%)	671.7 (6.8%)	655.2 (5.2%)	660.4 (5.0%)	649.0 (4.1%)	655.8 (6.4%)	655.8 (4.2%)	700.8 (4.6%)	704.7 (4.4%)
	-16.8 (-0.4%)		-20.9 (-4.5%)		-16.5 (-1.6%)		-11.4 (-0.9%)		0 (-2.2%)		3.9 (-0.2%)	

Notes. Significance codes: '**' < .01, '*' < .05, †: *p* < .10

Appendix J: Outputs from Mixed-Effects Models for the Familiarity Rating Task

1. The Effects of Participant Group and Stimulus Condition on Familiarity Rating 2 (“I Do Not Know the Meaning of This Word”)

Fixed effects				
Predictor	Estimate	*SE*	*z*	*p*
(Intercept)	-1.46	0.75	-1.94	.052
Condition	-0.64	0.21	-3.00	.002**
Group	-1.35	0.23	-5.90	< .001***
Condition:Group	-0.04	0.06	-0.67	.506
Random effects	Variance	SD		
Participant	1.32	1.15		
Item	8.41	2.90		

Notes. Formula (glmer): Familiarity ~ Condition * Group + (1 | Participant) + (1 | Item),
Significance codes: ‘***’ < .001, ‘**’ < .01

2. The Effects of Participant Group and Stimulus Condition on Familiarity Rating 2 with the /b/-/t/ or /d/-/t/ Contrast as the Reference Category for Stimulus Condition

Fixed effects				
Predictor	Estimate	*SE*	*z*	*p*
(Intercept)	-4.05	0.84	-4.82	< .001***
Condition1 (The /l/-/ɹ/ contrast)	1.93	0.94	2.05	.040*
Condition2 (The /f/-/p/ contrast)	1.83	0.96	1.92	.055
Condition4 (The homophone)	-0.09	0.98	-0.09	.930
Condition5 (The same initial)	-0.24	1.03	-0.23	.818
Group	-1.56	0.25	-6.16	< .001***
Condition1:Group	0.26	0.23	1.11	.270
Condition2:Group	-0.08	0.26	-0.32	.753
Condition3:Group	0.27	0.27	1.00	.314
Condition4:Group	-0.14	0.36	-0.40	.690
Random effects	Variance	SD		
Participant	1.33	1.15		
Item	8.48	2.91		

Notes. Formula (glmer): Familiarity ~ Condition * Group3 + (1 | Participant) + (1 | Item),
Significance codes: ‘***’ < .001, ‘**’ < .01, ‘*’ < .05

Appendix K: Output from Mixed-Effects Models for the Translation Task

1. The Effects of Participant Group and Stimulus Condition on Translation Accuracy

Fixed effects				
Predictor	Estimate	*SE*	*z*	*p*
(Intercept)	0.706	0.52	1.35	.178
Condition	0.614	0.16	3.85	< .001***
Group	0.934	0.16	5.71	< .001***
Condition:Group	-0.001	0.05	-0.02	.981
Random effects	Variance	SD		
Participant	0.45	0.67		
Item	4.27	2.07		

Notes. Formula (glmer): Accuracy ~ Condition * Group + (1 | Participant) + (1 | Item),
Significance codes: '***' < .001, '**' < .01

2. The Effects of Participant Group and Stimulus Condition on Translation Accuracy with the /b/-/t/ or /d/-/t/ Contrast as the Reference Category for Stimulus Condition

Fixed effects				
Predictor	Estimate	*SE*	*z*	*p*
(Intercept)	3.43	0.64	5.88	< .001***
Group	0.93	0.20	4.69	< .001***
Condition1 (The /l/-/ɹ/ contrast)	-1.80	0.70	-2.57	.010
Condition2 (The homophone)	-1.56	0.70	-2.23	.026
Condition3 (The /f/-/p/ contrast)	-1.57	0.70	-2.23	.026
Condition5 (The same initial)	0.64	0.78	0.82	.410
Condition1:Group	-0.04	0.21	-0.17	.863
Condition2:Group	-0.33	0.21	-0.17	.868
Condition3:Group	0.35	0.22	0.24	.809
Condition4:Group	-0.13	0.27	-0.30	.766
Random effects	Variance	SD		
Participant	0.45	0.67		
Item	4.22	2.05		

Notes. Formula (glmer): Accuracy ~ Condition * Group3 + (1 | Participant) + (1 | Item),
Significance codes: '***' < .001

Appendix L: Results of Exploratory Item-based Analyses for Determining the Reasons for Form Inhibition for L2 Speakers under the Same Initial Overlap Condition

- **The Patterns of Form Priming for Each Participant Group under the Same Initial Overlap Condition Depending on the Position of Different Letters (Analysis 1) and the Existence of a Shared Body (Analysis 2)**

	Analysis 1				Analysis 2			
	Pairs differing in a middle letter (*then-TEEN*) (*n* = 8)		Pairs differing in the last letter (*feed-FEET*) (*n* = 8)		Pairs sharing the body (*model-MOTEL*) (*n* = 6)		Pairs without a shared body (*black-BLOCK*) (*n* = 10)	
	Related	Unrelated	Related	Unrelated	Related	Unrelated	Related	Unrelated
Low-prof. L2 speakers	846.6	816.9	836.7	783.0	822.3	811.8	854.7	793.3
	-29.7†		-53.7		-10.5		-61.4	
Mid-prof. L2 speakers	763.3	737.5	736.8	740.6	776.7	722.0	733.4	749.5
	-25.8		3.8		-54.7		16.1	
High-prof. L2 speakers	825.1	810.3	835.8	847.9	818.7	836.7	833.8	823.6
	-14.8		12.1		18.0		-10.2	
L1 speakers	668.1	645.3	640.0	665.9	672.5	665.3	643.3	649.9
	-22.8		25.9		-7.2		6.6	

Notes. Significance codes: †: $p < .10$

References

Andrews, S. (1992). Frequency and neighborhood effects on lexical access: Lexical similarity or orthographic redundancy? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(2), 234–254.

Andrews, S. (1997). The role of orthographic similarity in lexical retrieval: resolving neighborhood conflicts. *Psychonomic Bulletin and Review*, *4*, 439–461.

Andrews, S., & Hersch, J. (2010). Lexical precision in skilled readers: Individual differences in masked neighbor priming. *Journal of Experimental Psychology: General*, *139*, 299–318.

Andrews, S. & Lo, S. (2012). Not all skilled readers have cracked the code: Individual differences in masked form priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*, 152–163.

Andrews, S., Veldre, A., & Clarke, I. E. (2020). Measuring lexical quality: the role of spelling ability. *Behavior Research Methods*, *52*(6), 2257–2282.

Barrios, S., Jiang, N., & Idsardi, W. J. (2016). Similarity in L2 phonology: Evidence from L1 Spanish late learners' perception and lexical representation of English vowel contrasts. *Second Language Research*, *32*(3), 367–395.

Baxter, P. A., Droop, W., Van Den Hurk, M. M., Bekkering, H., Dijkstra, A. F. J., & Léoné, F. T. M. (2021). Contrasting similar words facilitates second language vocabulary learning in children by sharpening lexical representations. *Frontiers in Psychology*, 12.

Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In M. J. Munro & O. -S. Bohn (Eds.), *Second language speech learning: The role of language experience in speech perception and production* (pp. 13-34). Amsterdam: John Benjamins.

Bijeljac-Babic, R., Biardeau, A., & Grainger, J. (1997). Masked orthographic priming in bilingual word recognition. *Memory & Cognition, 25,* 447–457.

Bordag, D., Gor, K., Opitz, A. (2021). Ontogenesis Model of the L2 Lexical Representation. *Bilingualism: Language and Cognition*, 1–17.

Borden, G., Gerber, A., & Milsark, G. (1983). Production and perception of the /r/-/l/ contrast in Korean adults learning English. *Language Learning*, *33*, 499-526

Bowers, J. S., Vigliocco, G., & Haan, R. (1998). Orthographic, phonological and articulatory contributions to masked letter and word priming. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 1705–1719.

Broersma, M. (2012). Increased lexical activation and reduced competition in second-language listening. *Language and Cognitive Processes*, *27*, 1205–1224.

Brown, J. D. (1980). Relative merits of four methods for scoring cloze tests. *The Modern Language Journal*, *64*(3), 311-317.

Brysbaert, M. (forthcoming). Word recognition II: Orthography-phonology. In M. Snowling, C. Hulme, & K. Nation (Eds.), Science of Reading: A Handbook (2nd ed.).

Brysbaert, M., Lagrou, E., & Stevens, M. (2017). Visual word recognition in a second language: A test of the lexical entrenchment hypothesis with lexical decision times. *Bilingualism: Language and Cognition*, *20*(3), 530–548.

Castles, A., Davis, C., & Letcher, T. (1999). Neighborhood effects on masked form-priming in developing readers. *Language and Cognitive Processes*, 14, 201–224.

Castles, A., Davis, C., Cavalot, P., & Forster, K. (2007). Tracking the acquisition of orthographic skills in developing readers: masked priming effects. *Journal of Experimental Child Psychology*, *97*(3), 165–182.

Colombo, L. (1986). Activation and inhibition with orthographically similar words. *Journal of Experimental Psychology: Human Perception and Performance, 12,* 226-234.

Coltheart, M. (2006). Dual route and connectionist models of reading: an overview. *London Review of Education*, *4*(1), 5–17.

Coltheart, M., Davelaar, E., Jonasson, J.T., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.)*, Attention and performance VI*, (pp. 535–555). Hillsdale, NJ: Erlbaum.

Cook, S. V., & Gor, K. (2015). Lexical access in L2: Representational deficit or processing constraint? *The Mental Lexicon*, *10*, 247–270.

Cook, S. V., Pandža, N. B., Lancaster, A. K., & Gor, K. (2016). Fuzzy nonnative phonolexical representations lead to fuzzy form-to-meaning mappings. *Frontiers in Psychology*, *7*, 1345-1345.

Cutler, A., Weber, A., Smits, R., & Cooper, N. (2004). Patterns of english phoneme confusions by native and non-native listeners. *The Journal of the Acoustical Society of America*, *116*(6), 3668–78.

Darcy, I., Dekydspotter, L., Sprouse, R. A., Glover, J., Kaden, C., McGuire, M., & Scott, J. HG. (2012). Direct mapping of acoustics to phonology: On the lexical encoding of front rounded vowels in L1 English-L2 French acquisition. *Second Language Research*, *28*, 5–40.

Davis, C. J. (2005). N-watch: a program for deriving neighborhood size and other psycholinguistic statistics. *Behavior Research Methods*, *37*(1), 65–70.

Davis, C., Castles, A., & Iakovidis, E. (1998). Masked homophone and pseudohomophone priming in adults and children. *Language and Cognitive Processes*, *13*, 625–651.

Davis, C. J., & Lupker, S. J. (2006). Masked inhibitory priming in English: Evidence for lexical inhibition. *Journal of Experimental Psychology: Human Perception and Performance*, *32*, 668–687.

Daidone, D., & Darcy, I. (2021). Vocabulary Size Is a Key Factor in Predicting Second Language Lexical Encoding Accuracy. *Frontiers in Psychology*, *12*.

Diependaele, K., Duñabeitia, J.A., Morris, J., & Keuleers, E. (2011). Fast morphological effects in first and second language word recognition. *Journal of Memory and Language*, *64*, 344–358.

De Moor, W., & Brysbaert, M. (2000). Neighborhood frequency effects when primes and targets have different lengths. *Psychological Research*, *63*, 159-162.

De Moor, W., Verguts, T., & Brysbaert, M. (2005). Testing the multiple in the multiple read-out model of visual word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *6*, 1502-1508.

De Moor, W., Van der Herten, L., & Verguts, T. (2007) Is masked neighbor priming inhibitory? Evidence using the incremental priming technique. *Experimental Psychology*, *54*(2), 113-119.

De Groot, A. M. B., & Keijzer, R. (2000). What is hard to learn is easy to forget: the roles of word concreteness, cognate status, and word frequency in foreign language vocabulary learning and forgetting. *Language Learning*, *50*(1), 1–56.

Dijkstra, T., & Van Heuven, W. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, *5*(3), 175-197.

Dufour, S., & Peereman, R. (2003a). Inhibitory priming effects in auditory word recognition: When the target's competitors conflict with the prime word. *Cognition*, *88*, B33–B44.

Dufour, S., & Peereman, R. (2003b). Lexical competition in phonological priming: Assessing the role of phonological match and mismatch lengths between primes and targets. *Memory & Cognition*, *31*(8), 1271–1283.

Eberhard, K. M. (1994). Phonological inhibition in auditory word recognition. In D. Dagenbach & T. H. Carr (Eds.), *Inhibitory processes in attention memory and language* (pp. 383–407). San Diego, CA: Academic Press.

Elgort, I. (2011). Deliberate learning and vocabulary acquisition in a second language. *Language Learning*, *61*, 367–413.

Elgort, I., & Warren, P. (2014). L2 vocabulary learning from reading: Explicit and tacit lexical knowledge and the role of learner and item variables. *Language Learning*, *64*, 365–414.

Ellis, N. C. (1994). Vocabulary acquisition: The implicit ins and outs of explicit cognitive mediation. In N. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 211–282). New York, NY: Academic Press.

Ellis, N. C. (1996). Sequencing in SLA: Phonological memory, chunking and points of order. *Studies in Second Language Acquisition*, *18*, 91–126.

Escudero, P., Hayes-Harb, R., & Mitterer, H. (2008). Novel second-language words and asymmetric lexical access. *Journal of Phonetics*, *36*(2), 345–360.

Ferrand, L., & Grainger, J. (1992) Phonology and orthography in visual word recognition: Evidence from masked nonword priming. *Quarterly Journal of Experimental Psychology*, *45A*, 353-372.

Ferrand, L., & Grainger, J. (1993). The time course of orthographic and phonological code activation in the early phases of visual word recognition. *Bulletin of the Psychonomic Society*, *31*(2), 119–122.

Flege, J. (2003). Assessing constraints on second-language segmental production and perception. In A. Meyer & N. Schiller (Eds.), *Phonetics and phonology in language comprehension and production: Differences and similarities* (pp. 319–355). Berlin: Mouton de Gruyter.

Flege, J. E., & MacKay, I. R. A. (2004). Perceiving vowels in a second language. *Studies in Second Language Acquisition*, 26(1), 1–34.

Forster, K. I. (1987). Form-priming with masked primes: The best-match hypothesis. In M. Coltheart (Ed.), *Attention & performance XII: The psychology of reading* (pp. 127-146). Hillsdale, NJ: Erlbaum.

Forster, K. I., & Davis, C. (1984). Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*, 680-698.

Forster, K. I., Davis, C., Schoknecht, C., & Carter, R. (1987). Masked priming with graphemically related forms: Repetition or partial activation? *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *39*(A), 211-251.

Forster, K. L, & Shen, D. (1996). No enemies in the neighborhood: Absence of inhibitory effects in lexical decision and semantic categorization. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *22*, 696-713.

Forster, K. I., & Veres, C. (1998). The prime lexicality effect: Form-priming as a function of prime awareness, lexical status, and discrimination difficulty. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 498–514.

Frisson, S., Bélanger, N. N., & Rayner, K. (2014). Phonological and orthographic overlap effects in fast and masked priming. *The Quarterly Journal of Experimental Psychology*, *67*(9), 1742-1767.

Frith, U. (1985). Beneath the surface of developmental dyslexia. In K. Patterson, J. Marshall, & M. Coltheart (Eds.), *Surface dyslexia* (pp. 301–330). Hillsdale, NJ: Erlbaum.

Gaskell, M. G., & Dumay, N. (2003). Lexical competition and the acquisition of novel words. *Cognition*, *89*, 105–132.

Gollan, T. H., Forster, K. I., & Frost, R. (1997). Translation priming with different scripts: Masked priming with cognates and noncognates in Hebrew-English bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*, 1122–1139.

Gor, K. (2018). Phonological priming and the role of phonology in nonnative word recognition. *Bilingualism: Language and Cognition*, *21*(3), 437–442.

Gor, K. (forthcoming). Word and multi-word processing: Memory-based and linguistic approaches. In A. Godfroid & H. Hopp (Eds.), *The Routledge Handbook of Second Language Acquisition and Psycholinguistics*.

Gor, K., & Cook, S. V. (2020). A mare in a pub? Nonnative facilitation in phonological priming. *Second Language Research*, *36*(1), 123-140.

Gor, K., Cook, S., Bordag, D., Chrabaszcz, A., & Opitz, A. (2021). *Fuzzy lexical representations in adult second language speakers*. Frontiers in Psychology, *12*.

Grainger, J., & Ferrand, L. (1994), Phonology and orthography in visual word recognition: Effects of masked homophone primes. *Journal of Memory & Language, 33*, 218-233.

Heyer, V. & Clahsen, H. (2015). Late bilinguals see a scan in scanner AND in scandal: Dissecting formal overlap from morphological priming in the processing of derived words. *Bilingualism: Language and Cognition*, *18*, 543–50.

Hamrick, P., & Pandža, N. (2020). Contributions of semantic and contextual diversity to the word frequency effect in L2 lexical access. *Canadian Journal of Experimental Psychology*, *74*, 25-34.

Humphreys, G. W., Evett, L. J., & Taylor, D. E. (1982). Automatic phonological priming in visual word recognition. *Memory & Cognition*, *10*, 576–590.

Jamieson, D. G., & Yu, K. (1996) Perception of English /r/ and /l/ speech contrasts by native Korean listeners with extensive English-language experience. *Proceeding of Fourth International Conference on Spoken Language Processing* (pp. 1453-1456), doi: 10.1109/ICSLP.1996.607889.

Jiang, N. (2000). Lexical representation and development in a second language. *Applied Linguistics*, *21*(1), 47-77.

Jiang, N. (2021). Orthographic friends and lexical strangers in the L2 lexicon. *Journal of Second Language Studies,* 202105.

Jiang, N., & Forster, K. I. (2001). Cross-language priming asymmetries in lexical decision and episodic recognition. *Journal of Memory and Language*, *44*, 32–51.

Jiang, N. & Pae, H. K. (2020). The pseudohomophone effect in visual word recognition in a second language. In Booth & Clenton (eds.), *First language influences on multilingual lexicons* (pp. 15-34). New York: Routledge.

Jiang, N., & Wu, X. (2022). Orthographic priming in second-language visual word recognition. *Language Learning*, (20220206).

Jiang, N., & Zhang, J. (2021). Form prominence in the L2 lexicon: Further evidence from word association. *Second Language Research*, *37*(1), 69–90.

Kim, S. Y. & Bolger. D. J. (2016). The role of subsyllabic units in the visual word recognition of Korean monosyllabic words: A masked priming study. *Journal of Cognitive Science*, 17(3), 343-359.

Kim, J., & Davis, C. (2003). Task effects in masked cross-script translation and phonological priming. *Journal of Memory and Language*, *49*, 484–499.

Kinoshita, S., & Norris, D. (2012a). Pseudohomophone priming in lexical decision is not fragile in a sparse lexical neighborhood. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(3), 764–75.

Kinoshita, S., & Norris, D. (2012). Task-dependent masked priming effects in visual word recognition. *Frontiers in Psychology*, 3.

Li, J., & Taft, M. (2020). The processing of English prefixed words by Chinese-English bilinguals. *Studies in Second Language Acquisition*, *42*, 239–249.

Li, J., Taft, M., & Xu, J. (2017). The processing of English derived words by Chinese-English bilinguals. *Language Learning*, *67*(4), 858–884.

Li, M., Jiang, N., & Gor, K. (2017). L1 and L2 processing of compound words: Evidence from masked priming experiments in English. *Bilingualism: Language and Cognition*, *20*, 384–402.

Linacre, J. M. (1997). KR-20 or Rasch reliability: Which tells the “Truth?”. *Rasch Measurement Transactions*, *11*(3), 580-1.

Lindsay, S., & Gaskell, M. G. (2010). A complementary systems account of word learning in L1 and L2. *Language Learning*, *60*, 45–63.

Llompart M (2021). Phonetic categorization ability and vocabulary size contribute to the encoding of difficult second language phonological contrasts into the lexicon. *Bilingualism: Language and Cognition*, *24*, 481–496.

Lukatela, G., & Turvey, M. T. (1994). Visual lexical access is initially phonological: 1. Evidence from associative priming by words, homophones, and pseudohomophones. *Journal of Experimental Psychology: General*, *123*, 107-128.

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, *102*, 419–457.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review*, *88*, 375–407.

Martin, K. I. & Ellis, N. C. (2012). The roles of phonological short-term memory and working memory in L2 grammar and vocabulary learning. *Studies in Second Language Acquisition*, *34*(3), 379–413.

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word recognition. *Cognition*, *25*, 71–102.

Martensen, H., Dijkstra, T., & Maris, E. (2005). A werd is not quite a word: On the role of sublexical phonological information in visual lexical decision. *Language and Cognitive Processes*, *20*(4), 513-552.

Meador, D., Flege, J. E., & MacKay, I. R. (2000). Factors affecting the recognition of words in a second language. *Bilingualism: Language & Cognition*, *3*(1), 55–67.

Mulder, K., Van Heuven, W. J. B., & Dijkstra, T. (2018). Revisiting the neighborhood: How L2 proficiency and neighborhood manipulation affect bilingual processing. *Frontiers in Psychology, 9,* 1860–1860.

Nakayama, M., Ida, K., & Lupker, S. J. (2016). Cross-script L2–L1 noncognate translation priming in lexical decision depends on L2 proficiency: Evidence from Japanese–English bilinguals. *Bilingualism: Language and Cognition*, *19*, 1–22.

Nakayama, M., & Lupker, S. J. (2018). Is there lexical competition in the recognition of L2 words for different-script bilinguals? An examination using masked priming with Japanese-English bilinguals. *Journal of Experimental Psychology: Human Perception and Performance*, *44*(8), 1168.

Nakayama, M., Sears, C. R., & Lupker, S. J. (2008). Masked priming with orthographic neighbors: a test of the lexical competition assumption. *Journal of Experimental Psychology: Human Perception and Performanc*e, *34*(5), 1236–1260.

Ota, M., Hartsuiker, R. J., & Haywood, S. L. (2009). The key to the rock: near-homophony in nonnative visual word recognition. *Cognition*, *111*(2), 263–269.

Ota, M., Hartsuiker, R., & Haywood, S. (2010). Is a fan always fun? phonological and orthographic effects in bilingual visual word recognition. *Language and Speech*, *53*(3), 383–403.

Pallier, C., Colomé, A., & Sebastián-Gallés, N. (2001). The influence of native-language phonology on lexical access: Exemplar-based versus abstract lexical entries. *Psychological Science*, *12*(6), 445-449.

Park, H., & De Jong, K. J. (2008). Perceptual category mapping between English and Korean prevocalic obstruents: Evidence from mapping effects in second language identification skills. *Journal of Phonetics*, *36*(4), 704–723.

Park, K. (2021). The effects of orthographic precision, frequency of and familiarity with L2 word primes on orthographic neighbor priming [Qualifying paper, University of Maryland].

Perfetti, C. A. (1992). The representation problem in reading acquisition. In P. B. Gough, L. C. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 145–174). Hillsdale, NJ: Lawrence Erlbaum Associates.

Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, *11*, 357–83.

Perfetti, C. A., & Hart, L. (2002). The lexical quality hypothesis. In L. Vehoeven. C. Elbro, & P. Reitsma (Eds.), *Precursors of functional literacy* (pp. 189–213). Amsterdam: John Benjamins.

Perfetti, C. A., & Tan, L. H. (1998). The time course of graphic, phonological, and semantic activation in Chinese character identification. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(1), 101–118.

Qiao, X., & Forster, K. I. (2013). Novel word lexicalization and the prime lexicality effect. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *39*(4), 1064–1074.

Qiao, X., & Forster, K. I. (2017). Is the L2 lexicon different from the L1 lexicon? Evidence from novel word lexicalization. *Cognition*, *158*, 147-152.

Rastle, K., & Brysbaert, M. (2006). Masked phonological priming effects in English: Are they real? Do they matter? *Cognitive Psychology*, *53*, 97-145.

Schmidt, A. M. (1996). Cross-language identification of consonants. Part1. Korean perception of English. *Journal of the Acoustical Society of America*, *99*, 3201–3211.

Segui, J., & Grainger, J. (1990). Priming word recognition with orthographic neighbors: Effects of relative prime–target frequency. *Journal of Experimental Psychology: Human Perception and Performance*, *16*, 65–76.

Shen, D., & Forster, K. I. (1999). Masked phonological priming in reading Chinese words depends on the task. *Language and Cognitive Processes*, *14*(5/6), 429–459.

Shipley, W. C. (1940). A self-administering scale for measuring intellectual impairment and deterioration. *The Journal of Psychology*, 9, 371–377.

Slowiaczek, L. M., & Hamburger, M. B. (1992). Prelexical facilitation and lexical interference in auditory word recognition. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *18*, 1239–1250.

Smith, C. (2001). L2 acquisition of English liquids: Evidence for production independent from perception. In X. Bonch-Bruevich, W. Crawford, J. Hellermann, C. Higgins, & H. Nguyen (Eds.), *Selected Proceedings of the 2000 Second Language Research Forum* (pp. 3-22). Somerville, MA: Cascadilla Press.

Shipley, W. C. (1940). A self-administering scale for measuring intellectual impairment and deterioration. *The Journal of Psychology*, *9*, 371–377.

Speciale, G., Ellis, N.C., & Bywater, T. (2004). Phonological sequence learning and short-term store capacity determine second language vocabulary acquisition. *Applied Psycholinguistics*, *25*, 293–321.

Thorn, A. S. C., & Gathercole, S. E. (2001). Language differences in verbal short-term memory do not exclusively originate in the process of subvocal rehearsal. *Psychonomic Bulletin & Review*, *8*(2), 357–364.

Van Heuven, W. J. B., Dijkstra, T., & Grainger, J. (1998). Orthographic neighborhood effects in bilingual word recognition. *Journal of Memory and Language*, *39*(3), 458–483.

Van Orden, G. C. (1987). A Rows is a Rose: Spelling, sound, and reading. *Memory and Cognition*, *15*, 181–198.

Wang, M., & Geva, E. (2003). Spelling acquisition of novel English phonemes in Chinese children. *Reading and Writing: An Interdisciplinary Journal*, *16*(4), 325–348.

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*(2), 129-149.

Weber, A., & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language*, *50*(1), 1–25.

Welcome, S. E., & Trammel, E. R. (2017). Individual differences in orthographic priming relate to phonological decoding skill in adults. *Cognitive Processing: International Quarterly of Cognitive Science*, *18*(2), 119–128.

Witzel, N. O., & Forster, K. I. (2012). How L2 words are stored: The episodic L2 Hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*, 1608–1621.