PHONOLOGICAL VARIATION AND L2 WORD LEARNING: THE ROLE OF ORTHOGRAPHY IN WORD RECOGNITION AND PRODUCTION

Chung-Lin Yang

Submitted to the faculty of the University Graduate School

in partial fulfillment of the requirements for the degree

Doctor of Philosophy

in the Department of Linguistics,

Indiana University

July 2015

Accepted by the Graduate Faculty, Indiana University, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Doctoral committee:

Isabelle Darcy, Ph.D.

Tessa Bent, Ph.D.

Kenneth de Jong, Ph.D.

Robert Port, Ph.D.

March 10, 2015

Acknowledgements

First and foremost, I want to wholeheartedly thank my family. I would not have accomplished so much without my parents, who encouraged and supported me to pursue my PhD. My mother, who passed away while I was pursuing my PhD, was always on my mind and with me. Also, the mental support from my wife, sisters and brother is beyond what I can describe here, especially my sister, Sandra, who always stepped in and gave me a hand whenever needed, and my dear wife, Zora, who is not only my beloved one but also my best friend and supporter.

It was my greatest honor to have been able to work with my tremendous advisors at IU. I can't find a word to express how grateful I am. This dissertation would not have been possible without my great supervisor, Professor Isabelle Darcy, who always supports and helps me in every aspect. I also want to thank my research committee members: Professor Bob Port, who inspired my interest in cognitive science and broadened my horizons on linguistics, Professors Kenneth de Jong and Professor Tessa Bent, who helped me dig into my data from different angles and provided me insightful feedback. Also thanks to Professor Robert Botne and Professor Stuart Davis for their advice and funding support for my study.

During the initial stages of designing the experiments, I had received a lot of invaluable feedback and suggestions from people at IU and audiences at various conferences. I especially want to thank Professor David Pisoni, who suggested me to add the production tasks to complement the findings from the recognition tasks, Professor Rachel Hayes-Harb (University of Utah), who gave me more ideas for future research as well as possible ways to improve my design, Danielle Daidone and Dr. Jeff Holiday, who both gave me very constructive suggestions on the experimental design, Ryan Lidster, who gave me very insightful feedback on the analysis of the production data , as well as other lab members at the Second Language Psycholinguistics Lab.

Lastly, I'd like to thank all my dear friends and participants. I could not have done so much without all your help and support.

Abstract

Chung-Lin Yang

PHONOLOGICAL VARIATION AND L2 WORD LEARNING: THE ROLE OF ORTHOGRAPHY IN WORD RECOGNITION AND PRODUCTION

Exposure to L2 orthography may help L2 learners distinguish a novel segmental (e.g., Escudero et al., 2008) and tonal contrast (Showalter & Hayes-Harb, 2013). However, little is known about whether learners can associate two phonetic variants with the same lexical entry when the orthographic form of the word is provided during learning. Thus, the goal of this dissertation is to investigate whether exposure to L2 orthography could help learners link two free variants with one lexical entry and whether Taiwanese and American participants will show different degrees of reliance on orthography in word learning. Four experiments were conducted in the current study. Experiment 1 examined whether orthography can help learners learn the [ɔ-u] free variation in an artificial language using a word learning paradigm modified from Hayes-Harb, Nicol, and Barker (2010). No effect of orthography on the learning of vocalic free variation was found in Experiment 1, and both L1 groups showed low detection of the vowel alternations. A follow-up picture naming task investigating learners' encoding and decoding of the words was conducted in Experiment 2. The results from Experiment 2 showed that orthography did help learners score higher in picture naming and also produce more new forms of the free variants. Experiment 3 and 4 paralleled 1 and 2, respectively, but the target alternation was consonantal (i.e., [p-b] or [t-d] counterbalanced). An effect of orthography on the learning of [p-b] free variation (but not [t-d]) was found in Experiment 3. Compared to Experiment 1, only Americans showed better detection of the consonantal alternations but not Taiwanese. Experiment 4, a picture-naming task, was a follow-up of Experiment 3. It was found that, unlike in Experiment 2,

orthography in this case did not always help learners score higher in production. In sum, mixed results regarding the effect of orthography on the learning of free variation were found. In general, the effect of orthography was stronger in production than in recognition. The findings reported in this dissertation add support to the role of orthographic information in lexical access.

Isabelle Darcy, Ph.D.

Tessa Bent, Ph.D.

Kenneth de Jong, Ph.D.

Robert Port, Ph.D.

Table of Contents

List of Tables

List of Figures

Chapter 1. Introduction

1.1 Word Learning and the Mental Lexicon

Word learning is a crucial step in language acquisition (McKeown & Curtis, 1987), which at least involves learning the pronunciation and meaning, and it is also a key process in building our 'dictionary in the brain' (i.e., mental lexicon, Aitchison, 2003: p.2). Some researchers view mental lexicon as a capacity for storing words (Weber & Scharenborg, 2012). However, relatively little is known about what cues or resources (e.g., auditory signal, orthographic forms, etc.) learners use to build up the lexicon in a second language (L2).

One of the most important tasks in L2 word learning is to learn the pronunciation of these lexical forms. A significant amount of research has shown that learning L2 pronunciation can be difficult because of various factors, such as the potential difference between the native language (L1) and L2 in terms of the sound inventory (e.g., Speech Learning Model, Flege, 1995; Perceptual Assimilation Model - L2, Best & Tyler, 2007) and the rules that govern sound sequencing (i.e., phonotactics) (e.g., de Jong & Park, 2012; Dupoux, Kakahi, Hirose, Pallier & Mehler, 1999). Thus, one challenge to successfully learn L2 pronunciation is to suppress the influence from L1.

The way pronunciation of L2 is learned can greatly vary from person to person. For example, some people may learn the pronunciation of a word by looking at the printed form and spelling out each sound, while some people may simply repeat an instructor's pronunciation without relying too much on the spellings. These two types of learning involve two different types of processing. For those who learn by spellings, they need to first convert each letter to the corresponding sound in the target L2, so they also need to process visual inputs from the written forms. But for those who repeat what an instructor says, they only need to process the auditory inputs, and to some extent the visual inputs from instructor's gestures. Thus, word learning, especially in educational contexts (regardless of L1 or L2), can involve both the recognition of printed words (visual word recognition) and spoken words (spoken word recognition). However, regardless of how pronunciation is learned, L2 learners do not always successfully achieve native-like pronunciation. First of all, knowing how to write and spell does not imply being able to accurately pronounce a word or achieve native-like pronunciation. Second, repeating an instructor's pronunciation can also be inaccurate, since phonological input is in most cases filtered through the native phonological system (e.g., Polivanov, 1931), which leads to adjustments both in perception and production (e.g., Flege, 1995). As mentioned earlier, pronunciation in L2 can be significantly influenced by L1 phonology, which includes the sound inventory, the restriction of sound combinations (phonotactics) and also the prosodic features.

One particular challenge in learning L2 pronunciation is the variability of pronunciation. As McQueen (2007) pointed out, a word could be stored in our memory as the "canonical form" of pronunciation (i.e., carefully pronounced form), or other forms (e.g., the casually produced variant). Imagine that if an L2 listener hears a word spoken in a casual way by another native speaker of the target L2, how could s/he associate this newly encountered variant with the word that s/he has learned before? One example would be the sequence "did you" [dɪdjuː] vs. "did ya" [dɪdʒə]. Or if the non-native listener learns the pronunciation of a word from a different variety of English, in which for instance there is no flapping (e.g., pronounce [t] instead of a flap in the word "water"), when s/he hears the flapped version of the word "water" later, would s/he be able to link this pronunciation variant to the word "water" or perceive it as a new word, if s/he has no prior exposure to and knowledge of the English flapping rule?

However, exposure to variability in speech can also be helpful in L2 phonological acquisition. Listeners exposed to experimental items produced by different speakers may be able to better distinguish a sound contrast that is not in his/her native language. Such research, the high-variability training studies (e.g., Bradlow et al, 1999), shows that listeners in fact store detailed cues (e.g., talker's voice, contextual effects) of the speech signal.

Nevertheless, one might wonder how L2 learners learn to link different pronunciation variants to the same lexical entry. One source could be the contexts, such as words embedded in sentences (i.e., semantic and syntactic contexts), which could possibly provide cues to help listeners recognize the pronunciation variant, so that the learner can understand the rest of the sentence. Another possibility for listeners to recognize words despite variability would be through exposure to the orthographic forms, which I will discuss in the next section.

1.2 The Role of Orthography in the Learning of L2 Words and Pronunciation

After literacy training, word learning can involve learning the orthography, especially through schooling as discussed earlier. In child language acquisition, as noted by Kouider and Dupoux (2001), hearing children first start perceiving the auditory input and gradually recognizing spoken words (regardless of first, second, or third language) before they are exposed to orthography and start developing literacy. The fact that the development of literacy is always preceded by the auditory input in first language acquisition can be considered a universal phenomenon, since even illiterate speakers of a language do not need to learn the orthography to acquire their L1. For L2 learners, especially adults, the situation is often very different: although it is also possible that "literate" learners can be immersed in an L2 environment to learn the target L2 without first being exposed to orthography, it is the case that in most formal educational contexts, learning the orthography of the L2 is an integral part of learning the language, to allow for both written and oral communication. Thus, both L1 and L2 speakers learn how to spell and write when they start to learn the vocabulary in the target language through formal education. For L2 learners, learning to write in the L2 orthography could be particularly challenging when the L2 uses a different orthography, such as English speakers learning Arabic or Chinese, or Chinese speakers learning English, since learning an entirely new orthography can be difficult. However, even when the same orthography is used, challenges and interferences from the L1 are to be expected: Indeed, just as it is the case for perception and production, the influence of the L1 extends to orthographic processing (e.g., Wang, Koda and Perfetti, 2003), as well. Therefore, another likely challenge is to learn the new letter-to-sound conversion rules when L1 and L2 share the same orthographic symbols. For example, the letter <H> in English corresponds to the phoneme /h/ but to the phoneme /n/ in Russian. In this case, English learners of Russian or Russian learners of English may find it difficult at first to suppress the influence of the L1.

Since orthography is so tightly associated with the learning of pronunciation after literacy training, this raises the major question of whether and how orthographic information is stored in the mental lexicon. Both Allport and Funnell (1981) and Fromkin (1987) suggested that orthographic information is also stored in the mental lexicon in addition to phonological and semantic information. Allport and Funnell (1981) also suggested a possible link between the phonological and orthographic forms in our mental lexicon. If orthographic information is stored in the mental lexicon, it is possible that orthographic information could influence how we recognize spoken words. This has indeed been shown: Ziegler and Ferrand (1998) found that when participants were asked to judge whether a stimulus was a word, words containing a rhyme that can be spelled in several ways in French elicited slower judgments than those with only one possible spelling. This finding may support that orthography plays a role in spoken word recognition and the formation of our mental lexicon. Sprenger-Charolles, Siegel, Béchennec, and Serniclaes (2003) also found that although phonological development at early stages is a predictor of later reading development, as literacy training progresses, the processing of phonological and orthographic information is interconnected in the process of visual word recognition and reading development.

In addition to the debate on when orthographic information comes into play during word recognition, another question is whether orthographic information could help L2 learners learn unfamiliar sounds that are not in their L1 (e.g., segmental or suprasegmental contrasts). Some learners might be able to make use of the spellings in the learning of pronunciation and also store both the orthographic and phonological forms of the words in their mental lexicon. For example, some studies have shown that exposure to L2 orthography may help L2 learners distinguish a novel phonemic contrast (e.g., Escudero, Hayes-Harb, & Mitterer, 2008), but there are also studies showing that exposure to orthography does not necessarily help for other contrasts (e.g., Hayes-Harb, Nicol, Barker, 2010; Escudero, Simon, Mulak, 2014). Generally, little is known about the extent to which orthographic information can help listeners recognize spoken words despite the variability in the speech signal. Specifically, can exposure to orthography help L2 learners recognize words pronounced in different ways? In other words, can orthographic forms help L2 learners associate pronunciation variants to the same lexical entry? This will be the overarching question for this dissertation.

1.3 Organization of this Dissertation

One central question in speech perception and spoken word recognition is what cues listeners have in perception and how listeners process the speech input and retrieve the corresponding lexical entry. The major goal of this dissertation is to investigate whether orthographic information plays a role in second-language learner's speech perception and word learning. In Chapter 2, I will first briefly discuss the various types of cues available to listeners when they recognize and learn words, including the role of orthography in speech perception and spoken word recognition. Then I will discuss the challenge for which orthography might prove useful to listeners – variation in speech. To close this chapter, I will discuss the models of lexical access and of the mental lexicon as well as how second-language learners encode novel phonemic contrasts in lexical representations. In Chapter 3, I will report findings from Experiment 1, a word learning experiment involving vocalic free variation. Chapter 4 reports the findings of a picture-naming task (Experiment 2), a follow-up experiment of Experiment 1. Chapter 5 reports findings from Experiment 3, which is a replication of Experiment 1 with new items involving consonantal free variation to investigate the effect of orthography on the learning of consonantal free variation. Chapter 6 is again a picture-naming task (Experiment 4), which is a follow-up of Experiment 3 to examine participants' spontaneous production of the words they learned. Finally, in Chapter 7, I will discuss the findings from all the four experiments, and offer some conclusions as well as outline future directions.

Chapter 2. Background

2.1 Speech Perception, Spoken Word Recognition and Lexical Representations

2.1.1 Perceptual cues for listeners

Listeners may exploit a wide range of cues when perceiving speech. Some researchers argue that perceiving speech is in fact perceiving the "intended" articulatory gestures, such as the motor theory of speech perception proposed by Liberman at al. (1967) (see also Liberman & Mattingly, 1985). McQueen (2007) indicated that the information listeners may use in spoken word recognition includes both segmental and suprasegmental cues. However, what listeners actually perceive is beyond this binary classification. During the process of spoken word recognition, there is an enormous amount of information in the signal that listeners can use to recognize spoken words given that speech is a continuous and multidimensional, rather than discrete, stream of signals. Although the speech signal is rich (e.g., Port, 2010; Palmeri, Goldinger, & Pisoni, 1993), not all the cues are relevant in the recognition process, and likewise listeners might not use all the cues to recognize spoken words. Instead, the cues that are relevant for listeners can be *language-specific*, that is, they can differ according to the phonology of listeners' native language.

For example, when recognizing two words differing only in one sound where one starts with a [p] and another starts with a [p^h], the voice onset time (VOT), measured from the stop release till the onset of voicing (Reetz & Jongman, 2011), is a feature that listeners may use to distinguish one from another if this contrast exists in his/her native language. Furthermore, instead of a binary classification based on aspiration (e.g., aspirated vs. unaspirated), VOT is a continuum such that languages can have different length of VOT to contrast meanings. For example, Thai has a three-way VOT contrast: voiced (negative VOT), voiceless unaspirated (short lag) and voiceless aspirated stops (long lag) (Ladefoged, 2004; Lisker & Abramson, 1963). Listeners from a language where there is only a two-way contrast may find it difficult to differentiate a three-way contrast. For example, according to Lisker and Abramson (1963), Dutch has a two-way VOT contrast (/b/-/p/). Thus, to Dutch listeners, [p] with longer aspiration (i.e., [p^h]) might still be considered as a realization of the /p/ phoneme (i.e., an allophone) or even considered as a sound produced by non-native speakers. One can imagine that Dutch listeners might have difficulties in differentiating the three-way contrast in Thai. The VOT continuum is one example of the temporal cues, which are another important source of information for listeners to recognize spoken words from a continuous speech stream.

Temporal cues are useful for listeners to discriminate not only long vs short-lag stop consonants but also vowels. Non-native listeners whose native language (L1) does not have a contrast of vowel length may find it difficult to distinguish short from long vowels. For example, distinguishing short from long vowels in Japanese can be challenging for French native speakers, because French does not have vowel length contrast (Dupoux et al., 1999). While in other languages where there is a tense-lax distinction, both the temporal and spectral cues are crucial for listeners in perception. For instance, the difference between English [e^I] and [ɛ] not only lies in the temporal but also the spectral difference, with [e^I] having a lower first formant frequency (F1) and longer duration, and [ɛ] having a higher F1 and shorter duration.

Since the speech signal is dynamic in nature, the articulation of a preceding or following sound will also affect the articulation of the sound in between. The seminal work by Öhman (1966) showed the coarticulatory effect in the VCV contexts with different consonants. Formant transitions have also been found to be important cues to spoken word recognition (e.g., Warren and Marslen-Wilson, 1987; Strange, Jenkins and Johnson, 1983). Native listeners can even use formant transitions and temporal cues to identify syllables without the syllable nuclei (Strange et al., 1983). In Strange et al.'s study, naturally produced syllables were edited to create "silent-center syllables", which were syllables where the syllable nucleus was removed. Their results showed that native listeners could identify 94% of the silent-center syllables overall, suggesting that the formant transitions into and out of the syllable nucleus may provide enough cues for native listeners to recognize spoken words.

The discussion above provides only some examples of temporal cues. Although it is possible to distinguish two sounds only by the temporal cues (e.g., long- vs. short-lag stops), the spectral properties of speech sounds could vary along with the temporal domain. Even though the variation could be minimal across different tokens of the same sound, the variation in spectral properties can be intrinsic in natural productions. Aside from temporal cues, listeners can also distinguish two sounds by spectral cues, such as different vowel qualities. Other spectral cues include changes in amplitude and fundamental frequency (F0). Changes in F0 can create differences in pitch. In tonal languages, the difference in pitch height is used to contrast meanings, so the pitch difference is one spectral cue that listeners of tonal languages use to recognize spoken words. For example, in Mandarin, tone is a crucial cue for listeners to discriminate words with the identical segmental structure (e.g., [ma1] 'mother' vs. [ma3] 'horse').

In addition to various cues discussed above, listeners have been shown also to exploit non-acoustic information, such as visual cues and orthography. This suggests that information from various sources is integrated during speech perception and word recognition. One example of the visual influence on speech perception is the well-known McGurk Effect (McGurk & Macdonald, 1976) in which listeners reported hearing [da] while in fact listening to [ba] and seeing the lip movements of [ga]. There are also studies showing the possible benefit of visual input in speech perception. For example, Kerzel and Bekkering (2000) found that visible gestures facilitates perception, so listeners would respond faster when they saw a mouth gesture that matched the target sound. Massaro and Jesse (2007) and Sumby and Pollack (1954) also found that listeners may benefit from visual speech information in auditory word recognition. Calvert et al. (1997) reported that attending to silent speech will activate auditory cortical areas that are mainly responsible for processing auditory signals.

2.1.2 Challenges in word recognition – Variations in speech

One of the characteristics of speech communication is that no one single word or sound produced by the same talker will always be the same in terms of acoustic properties, so even reading a word list repetitively, the acoustic properties of each repetition will never be the same. Variation in speech also needs to be taken into account in the models of spoken word recognition (Luce & McLennan, 2008). In addition to within-talker variation, there is also a large amount of inter-speaker variation, such as dialectal variation (Clopper & Pisoni, 2008). Variation in speech is one challenge that listeners have to overcome to successfully recognize spoken words (Connine & Pinnow, 2006; Kraljic, Brennan, & Samuel, 2008). A small change in frequencies, aspiration, closure, release, voice quality, duration or amplitude can all create different extent of variation. Some variation may cause meaning changes, which are dependent upon language-specific phonology. For instance, voice quality can be contrastive in some languages (e.g., breathy voice in Tsonga, a Bantu language, and Marathi, an Indo-Aryan language) (Ladefoged, 2004), while such variation in voice quality does not contrast meanings in other languages (e.g., English, German, Mandarin, and Japanese).

How can a listener recognize tokens of the same word if there is always variation in the signal? One proposed account is that listeners may form prototypes for the similar speech tokens, and the development of prototypes of speech sounds emerges as early as infancy (Kuhl, 1991; Kuhl et al., 1992). The prototypes of speech sounds are important for listeners to recognize spoken words despite high variability of the speech signal. In addition, it should be noted that listeners in a speech community may have the same prototype of a sound but not necessarily produce exactly the same token. When a listener hears a token from a prototype, the goodness of fit or the distance between the token and the prototype will determine whether the perceived token is an instance of the prototype. When a perceived token is too “far” away from the prototype shared by the speech community, it could be a mispronounced target or produced by a speaker outside of the speech community, which could be from either a different variety of the same language or a different language. Thus, it is conceivable that non-native listeners may form a prototype of the same category different than native listeners, especially when the non-native listeners’ L1 does not have a similar category to that in L2. Best (1994) further proposed the Perceptual Assimilation Model (PAM) to account for the perception of non-native speech sounds. Under this framework, there are several possible category assimilation scenarios, one of which is the “single-category assimilation” where listeners perceive two non-native sounds as the same category in their L1. If we use the prototype model as an analogy, a token in prototype A, for example, might be mistakenly categorized as a token in prototype B by non-native listeners due to the influence from his/her native language. Hence, accurately categorizing non-native speech sounds can be challenging for non-native listeners, especially when encountering new tokens produced by different talkers.

Recognizing new tokens produced by different talkers could be challenging to not only non-native but also native listeners especially if the talker is from a different dialectal region. As mentioned earlier, the speech signals contain a large amount of phonetic detail. It is also closely associated with individual talker's physiological properties involved in speech production, which result in, for example, talker-specific voice and pitch range, constituting the "indexical properties of speech" (e.g., Pisoni, 1993). The indexical properties are also sources that cause variation in speech. The importance of the indexical properties of speech in speech perception and spoken word recognition has been demonstrated in speech training studies. It has also been found that exposure to variation may be beneficial for second-language learners in acquiring an unfamiliar contrast. Pisoni and colleagues (Bradlow, Akahane-Yamada, Pisoni, & Tohkura, 1999; Bradlow, Pisoni, Akahane-Yamada, & Tohkura, 1997; Lively, Logan, & Pisoni, 1993; Logan, Lively, & Pisoni, 1991; Pisoni & Lively, 1995) developed a training technique that incorporated high variability among talkers and also phonetic contexts to train Japanese learners of English to identify English /r/ and /l/. What they found was that Japanese learners of English who received high-variability training were better at generalizing the acquired contrast /r/ and /l/ to tokens produced by new talkers as well as in other phonetic contexts even after an extended period of time. Additionally, as found in Bradlow et al. (1997), Japanese learners of English not only benefited from high-variability training in their perception but also production. Although there was considerable between-subject variation, generally they still showed improvements in production and perception. However, a recent study by Perrachione, Lee, Ha, and Wong (2011) showed that high-variability training may not be beneficial to learners unless they have good perceptual abilities. They further argued that any training paradigm should take individual difference into account, since each person may possess different perceptual and cognitive abilities.

From perceptual training studies, we know that exposure to variation (i.e., variable context, voice, etc.) may be beneficial for learning an unfamiliar phonemic contrast. However, there are also variations that occur at the lexical level where a sound is dropped or changed to another sound, or an additional sound is inserted, which does not cause meaning change. An example of such a variation is the French word "cherry tree", which can be pronounced as either [səʀizje] or [sʀizje] (Bürki, Alario, & Frauenfelder, 2011) regardless of the phonological context (i.e., free variation). Although the two variants can be produced interchangeably by native speakers of French, will French learners, especially beginners, be able to associate both variants with the same meaning "cherry tree"? Coenen, Zwitserlood, and Bölte (2001) argued that variation in speech signal that causes "mismatch between input and lexical representation" is not tolerable except for "phonologically regular variation". Put differently, words that contain context-free variation will be harder for listeners to recognize than those that contain phonologically conditioned variation. Even infants were also shown to be able to react to variation in speech signal. For example, Swingley and Aslin (2000) found that 18-to-23-month-old toddlers could not recognize familiar words very well when words were slightly mispronounced (e.g., baby – vaby), which in this case is context-free variation. The mapping from the acoustic signal to the lexical representation has been one of the critical issues in understanding the process of spoken word recognition in both L1 and L2.

2.1.3 What is lexical representation?

Elman (2004) and Phillips (1999) indicated that each lexical entry contains the phonological, syntactic and semantic information about a word. All these pieces of information are necessary for successfully producing phrases and sentences. One piece of evidence supporting the role of lexical representations in the mental lexicon comes from the behavioral difference between words and nonwords. For example, Vitevitch and Luce (1998) found that in a speeded auditory shadowing task participants' performance was significantly different in real words and nonwords. For high-phonotactic probability/high-neighborhood density stimuli, nonwords were repeated more quickly and accurately, whereas real words were repeated more slowly. The reason accounting for this difference is that high-probability phonotactics usually occur in high frequency words and thus result in intense competition. However, nonwords do not have links to lexical representations and thus high probability sound sequences will yield faster responses. Also, Marslen-Wilson and Warren (1994) found that cross-spliced stimuli (mismatched coarticulatory information) slowed down processing of real words but not nonwords. They attributed this difference between words and nonwords to the lack of lexical representations for nonwords.

Given the findings presented here, we might postulate that there is a distinct mental lexicon from which real words are retrieved in the process of recognition, while nonwords do not have lexical entries in the mental lexicon and hence a behavioral difference between real and nonwords is observed in previous studies.

An even more complicated and debatable issue is whether bilinguals have two or a unified mental lexicon. As Kroll and Sunderman (2003) indicated, there still exists large disagreement as to whether for bilinguals, words in one language have a different lexical representation from the words in another language, but they also noted that words in different languages may still share the same concept node in the higher level.

As mentioned in Chapter 1, Allport and Funnell (1981) proposed that there are three basic components in the mental lexicon, namely, the cognitive code, the orthographic code and the phonological code, and there are interactive links among these three codes. Alternatively, Fromkin (1987) proposed a model of mental lexicon that incorporates three different subparts that are slightly different from the one proposed by Allport and Funnel – phonological lexicon, orthographic lexicon, and semantic lexicon. Additionally, she also proposed a system called "grapheme-phoneme rules" that mediates between phonological and orthographic lexicon. She reported data from individuals diagnosed with various types of dyslexia – those who had significant difficulties in reading. Her data showed that a deep dyslexic (i.e., a dyslexic who is unable to read out the letters) can produce a word only if it is first accessed through its meaning, and there seems to be no link between the phonological and orthographic code in this type of dyslexia. Hence, the separate modality for written and spoken words in the mental lexicon seems to be well established given the behavioral evidence reported here and the models proposed by Fromkin (1987) and Allport and Funnell (1981).

2.1.4 Mediated vs. Direct lexical access

The question of how speech signals are mapped to lexical representations to retrieve the corresponding lexical entry in the mental lexicon has been a central debate. Essentially there are two major views: direct mapping and mediated mapping (see McLennan, Luce & Charles-Luce, 2003 for a review). According to McLennan et al. (2003), the two frameworks differ in the levels of processing between speech signals and the lexical representations, where the mediated mapping presumes that the signals will be analyzed into features and abstract representations before being mapped onto the lexical forms (e.g., Lahiri, 1991; Norris, 1994; McClelland & Elman, 1986), whereas there are no mediated levels between the acoustic signals and lexical representations in the direct access framework (e.g., the "Lexical Access from Spectra" (LAFS) proposed by Klatt (1979)).

Although studies have shown that listeners store acoustic details including indexical properties in the lexical representation (e.g., Pisoni, 1993), many researchers have argued for prelexical phonological abstraction during lexical access, which is another example of mediated mapping (Dupoux, Pallier, Kakehi, & Mehler, 2001; McQueen, Cutler, & Norris, 2006; Scharenborg, Norris, ten Bosch, & McQueen, 2005).

In an attempt to determine whether lexical access would be mediated by an abstract underlying representation, McLennan et al. (2003) conducted a series of priming experiments in which the primes were either casually or carefully articulated intervocalic alveolar (i.e., /t-d/) or non-alveolar contrasts (i.e., /p-b/) where the alveolar ones were flapped and the non-alveolar ones were not when "casually" produced. They found that the primes containing flaps (e.g., [æɾəm]) could prime the targets without flaps (i.e., the canonical form [ædəm]), and the reverse was also true: the primes without flaps (i.e., the canonical form) could prime the targets with flaps. This finding supports the view that the lexical access can be mediated by a sublexical level from the speech signal. Nonetheless, the results from the non-alveolar consonants yielded a different finding: when the speaking style of the primes and target did not match (e.g., casual vs. careful or the reverse), the priming effect was smaller than that with matched speaking styles. This finding, however, did not support the mediated lexical access model. In the case of mediated access, the casually spoken prime should have also primed the carefully spoken targets, and vice versa, because the casually spoken prime will first map onto a sublexical representation that matches the canonical form.

Another study conducted by Kemps, Ernestus, Schreuder, and Baayen (2004) also finds support for a prelexical representation of lexical access. In Dutch, the phoneme [l] in the suffix "-(e)lijk [(ə)lək] (-ly)" can be dropped in running speech, which results in the reduced form of the suffix. One example they gave is the word "eigenlijk" (actually) [εiχələk], which can be reduced to [εiχək]. In other words, [εiχələk] and [εiχək] are in free variation. Using a phoneme monitoring paradigm in which listeners had to monitor the occurrence of [l], they investigated whether listeners could "restore" the reduced form – whether they would answer "yes" (the presence of [l]) to the reduced form. Their results revealed that when the targets were presented in isolation, listeners answered significantly more "no" than "yes" to the reduced forms. However, when the reduced forms were presented in contexts, there were a high percentage of false alarms – listeners answered "yes" to the reduced forms embedded in contexts. That is, they "restored" the reduced form to the canonical one. They concluded that there should be only one lexical representation for production variants. In addition, if the acoustic signal is directly mapped onto the lexical representation, the phoneme restoration process would not have been observed in the study. The authors also attributed the observed effect to orthography – listeners may resort to the orthographic representation of the canonical form when they monitored the target phoneme [l]. Thus, their results argue for the presence of a mediated level from the signal to the lexical representation.

2.1.5 Lexical representations and novel phonemic contrasts

Another central question regarding lexical representation and mental lexicon is whether second language learners are able to encode novel phonemic contrasts in the lexical representation. Darcy et al. (2012) found that advanced learners of French could not reach native-like performance on the categorization of the novel phonemic contrasts /œ/-/ɔ/ and /y/-/u/, but were able to build lexical representations that contain the novel contrasts – the previous presentation of /u/ did not significantly facilitate the lexical decision of words containing /y/, and vice versa, suggesting that in advanced learners' lexical representations, the words containing /y/ are different from those containing /u/, and the reverse is also true. But advanced learners' performance on an ABX categorization task was significantly poorer than native speakers', suggesting a difficulty in categorizing the novel contrast. Overall, their findings suggest that the ability to build novel lexical representations does not guarantee nor rely on native-like categorization of the novel phonemic contrast.

Similar findings can also be found in Cutler, Weber and Otake (2007). Using an eye-tracking paradigm, they found that Japanese participants looked at the picture of "locker" longer when the target picture was "rocket". However, when the target is "locker", they tended not to look at "rocket". Their findings suggest that: 1.) in listeners' lexical representations there is still a distinction between the contrast that they could not well distinguish in a pure auditory categorization task; and 2.) the asymmetry between the results from /r/ and /l/ may be accounted for by the fact that /l/ is the dominant (i.e., more familiar) category, which is consistent with the category in listeners' L1. Their findings suggest that listeners may be able to encode the phonological contrast in their lexical representation (although not as well as native speakers), even though they may still merge the two categories into one at the perceptual level.

In another similar study on the encoding of /r/ and /l/ in the lexical representation by Japanese, Ota, Hartsuiker, and Haywood (2009) used a semantic-relatedness judgment task (Luo, Johnson, & Gallo, 1998) in which pairs of words that were semantically related (e.g., KEY-LOCK) or unrelated (e.g., KEY-ROCK) were visually presented to participants. The crucial manipulation was the contrast /r/-/l/ in both pairs. As predicted, Japanese participants produced significantly more false alarms (~24%) than native English speakers (~4%) on the pairs contrasting in /r/-/l/. This result is striking because even though the participants saw the spellings, they still tended to accept "ROCK" as a semantically related word to "KEY", suggesting that they still activated /l/ (for "LOCK") when seeing <R> in "ROCK". In other words, they could not clearly differentiate /r/ from /l/ in their lexical representations. English native speakers, on the other hand, made only about 5% errors on the /r/-/l/ condition. Although the experiment in this study was purely visual, it provides insights for the role of non-native contrasts in lexical access from visual word recognition. It is noteworthy that in purely auditory tasks, such as Darcy et al.'s study (2012), non-native listeners may be able to build two separate lexical representations for a pair of words containing a non-native contrast even if they may not be able to categorize the contrast like native speakers do. Although the findings from visual and auditory tasks might not be comparable, it is also possible that listeners could activate the orthographic forms in an auditory task, which might benefit them in differentiating the words with difficult contrasts. Likewise, when recognizing words visually, the phonological information was activated, which in turn confused Japanese participants due to the lack of /r/-/l/ contrast in their L1.

The issue of what forms are stored in the mental lexicon still remains debatable (e.g., Weber & Scharenborg, 2012). Note that in the mediated access models, the levels between the acoustics and lexical representations consist of features and abstract phonological units, but is it possible that orthographic forms also play a role in auditory lexical access? If so, when and how are they used? In addition, after literacy training, is it possible that literacy changes the way words are stored in our mental lexicon? In particular, is there any difference between speakers of different writing systems in terms of how they store words in memory?

In section 2.2, I will discuss the differences between the major categories of writing systems and how they affect word recognition and lexical access. Following this line of argument, I will further discuss the effect of orthography on second language phonological acquisition and the background for the current study.

2.2 Orthographic and Speech Processing

2.2.1 Orthographic processing

2.2.1.1 Orthographic depth and L1-L2 orthography mapping

Learning new words, especially through formal schooling, involves learning their pronunciations, meanings, spellings and their syntactic functions. Learning the pronunciation with spelling could be challenging if the grapheme-to-phoneme correspondence in that particular L2 is different from that in L1, especially when L1 and L2 share the same orthography (e.g., French and English) or the orthographic depth between L1 and L2 is mismatched. The Orthographic Depth Hypothesis (ODH), according to Katz and Frost (1992: 150) states that

> "Shallow orthographies [i.e., consistent grapheme-to-phoneme correspondence] are more easily able to support a word recognition process that involves the language's phonology. In contrast, deep orthographies [i.e., inconsistent grapheme-to-phoneme correspondence] encourage a reader to process printed words by referring to their morphology via the printed word's visual-orthographic structure."

Based on this hypothesis, English, in their definition, is a deep orthography where one grapheme may correspond to several phonemes (e.g., <u> may correspond to [u], [ʌ] or [ʊ]). Spanish and German, for example, are shallow orthographies where grapheme-to-phoneme correspondence is more consistent than that in English.

Consider a Spanish learner of English learning to spell English words, s/he may find it challenging since the English grapheme-to-phoneme correspondence is less consistent than Spanish, which may hinder his/her learning outcome. For example, the letter <u> only corresponds to [u] in Spanish but can correspond to [u], [ʌ] or [ʊ] in English as mentioned above, so beginning Spanish learners of English would need to learn this new grapheme-to-phoneme correspondence. However, when an English learner of Spanish is learning the Spanish spellings, s/he may find it less challenging since the grapheme-to-phoneme correspondence is more consistent than that in English. In general, the inconsistent grapheme-to-phoneme correspondence will be harder for learners (both L1 and L2) to learn than the consistent one.

So far we have only discussed the orthographic mapping between two alphabetic writing systems, a similar question arises for other types of writing systems, such as when one language is alphabetic and the other is not, such as the logographic writing system used by Chinese. It is very common for user of a logographic language (e.g., Chinese) to learn an alphabetic language (e.g., English), so the issue of the mapping between an alphabetic and logographic writing must be addressed in cross-linguistic visual word recognition. Also, this issue should be addressed in spoken word recognition, since it has been found that orthography can influence speech perception and spoken word recognition (e.g., Ziegler & Ferrand, 1998).

Depending on the linguistic units that the orthographic form represents, the major writing systems in the world can be categorized into either *morphographic* writing (i.e., graphemes correspond to morphemes), including logographic, or *phonographic* writing (i.e., graphemes correspond to sounds), including syllabic (e.g., Japanese) and alphabetic (e.g., English) (Rogers, 2005: 13-15). Since morphographic and phonographic writing systems differ in what the graphemes correspond to, the users of these two types of writing systems might also process the orthographic and phonological information of the written words in different ways. Before further addressing the processing difference and the mapping between two distinct writing systems, the following section will first compare the difference between an alphabet (a type of phonographic writing) and a logograph (a type of morphographic) and discuss the characteristics of the logographic writing system.

2.2.1.2 Alphabetic vs. logographic writing system

In an alphabetic writing system, the spelling of a word usually provides a cue for pronunciation. One of the major differences between an alphabetic and logographic writing system is that there is usually not a systematic correspondence between a logograph and its pronunciation. As Steffensen, Goetz, and Cheng (1999) indicated, Chinese characters are logographic, containing a great amount of graphic information, but the grapheme-to-phoneme correspondence in Chinese characters is not as direct as in English (Wang, Koda, & Perfetti, 2003). However, DeFrancis (1989) argued that Chinese should not be classified as a logographic writing system but a morpho-syllabic one. In other words, each word (or morpheme) is a syllable. He further pointed out that a large portion of Chinese characters include a phonetic element (i.e., a subpart of the character that corresponds to a phoneme or syllable), which serves as a cue in recognizing the pronunciation. However, the fact is that even these phonetic elements are derived from logographs, and these elements can also be independent characters (Koda, 1989). For example, the character '山' (*shan* 'mountain'), which can also be a phonetic element in other characters, is derived from the shape of a mountain or hill to indicate the meaning of the character. Furthermore, unlike an alphabet writing system like English, it is impossible to combine any two phonetic elements in Chinese to create a new word. Zhou (1978) pointed out that using phonetic elements to predict pronunciation is not very reliable in Chinese. Consequently, to argue that Chinese writing is phonologically based may need more justification from the perspective of language processing. In a study on Chinese word recognition done by Seidenberg (1985), he found that the regularity effect (faster character naming for characters with the pronunciation being the same as the phonetic element) was larger in low-frequency words. When Chinese speakers are unsure about the pronunciation of a low-frequency character, they might pronounce the character based on the phonetic element (if any). However, the pronunciation of the character is not necessarily the same as that of the phonetic element. Thus, pronouncing the character solely based on the phonetic element could be unreliable as Zhou (1978) pointed out.

The processing of Chinese characters may also be explained by the dual-coding theory of memory and cognition proposed by Paivio (1986, 2007), in which there are two routes of processing – verbal and nonverbal coding. The two processing systems are independent of each other but interconnected, and both systems can operate simultaneously. For example, one can describe an image or object while seeing it. In his model, both spoken and visual word recognition are processes of verbal coding, while object and image recognition are processes of nonverbal coding. However, he also noted that for logographic readers like Chinese, recognizing a character might also involve imagery representation given the pictorial nature of logographs. When a Chinese speaker reads a character, the imagery representation will be activated through the interconnection with the verbal coding system.

Koda (1989) classified the major writing systems in the world into two categories: phonographic and morphographic, based on whether phonological structure of the word could be manifested on the orthographic form. In her definition, phonographic writing, including both alphabetic and syllabic writing systems, is phonologically based and the phonological code is visually accessible, whereas morphographic writing, including logographic writing (e.g., Chinese, Japanese Kanji and Korean Hanja), is based on morphemes and the phonological code may not be accessed directly from the orthographic forms. She further argued against the use of 'radicals' (phonetic elements) as a representation of phonological information. Several experimental studies that will be discussed later will provide some supportive evidence for her argument on the classification of Chinese writing.

Note that in an alphabetic writing system, letters are not necessarily always serially ordered. For example, in the Korean alphabet, Hangul letters are grouped into syllable blocks (see Figure 2.1 below), instead of being written sequentially like the letters of the Latin alphabet as in English and most of other alphabetic languages (Cho and Chen, 1999).

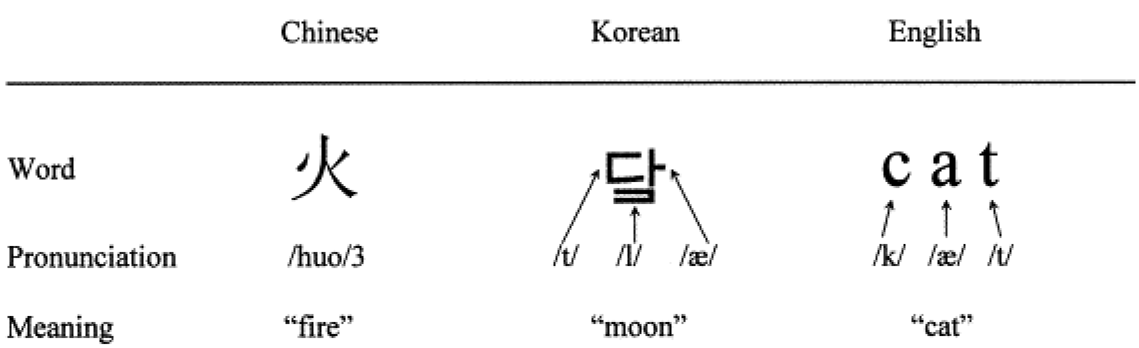


Figure 2.1. A comparison of Chinese, Korean and English writing systems. Reprinted from "Alphabetic and nonalphabetic L1 effects in English word identification: a comparison of Korean and Chinese English L2 learners" by M. Wang, K. Koda and C. Perfetti, 2003, *Cognition, 87,* p. 130. Copyright 2002 by Elsevier Science. Reprinted with permission.

In addition, according to Cho and Chen (1999), Hangul may be processed differently from Hanja (e.g., 門 'door'), which is borrowed from Chinese characters. They noted that phonological information is activated in the process of recognition of Hangul, whereas phonological mediation is not found when processing Hanja. They also found that Koreans can switch processing strategies while reading a mixed text containing both Hangul and Hanja, which is also a piece of evidence that processing logographs can be different from alphabets.

In sum, one major difference between Chinese and an alphabetic language is that Chinese logographs are associated with meaning more closely than alphabetic words, as stated in the Direct-image Hypothesis proposed by Aaronson and Ferres (1986). The hypothesis states that a single Chinese character carries more imagery information that comes from the logographs than an alphabetic word. On the other hand, the correspondence between orthographic forms and sounds in alphabetic languages is more directly linked than that in logographic languages. Thus, compared to an unknown English word, it would be easier to guess what meaning an unknown Chinese character might be related to. Based on the Orthographic Depth Hypothesis, we may categorize English as a shallow orthography compared to Chinese, which can be categorized as a deep orthography. However, note that orthographic depth should be a relative rather than absolute distinction. Although compared to other alphabetic languages, such as German and Spanish, English is considered as a deep orthography. Chinese, given its logographic nature, has phoneme-grapheme correspondence that is more "opaque" than English (see Table 2.1 below for a comparison of three languages on a hierarchy of orthographic depth). Also, the orthographic depth is determined by the "phoneme"-grapheme correspondence, suggesting that one grapheme can still correspond to different "allophones". In that sense, even in a very shallow orthography, there can be still some less than transparent form-to-sound mapping.

Table 2.1. A sample hierarchy of orthographic depth ordered from the deepest (left) to the shallowest orthography (right)

Languages	Chinese	English	Spanish
Grapheme-to-phoneme correspondence	Weak (e.g., ‘多’ (many) is [t^{w}o] but the character ‘侈’ (a last name) containing ‘多’ is [t^{h}ɔŋ])	Weak but stronger than Chinese (e.g., <u> → [u][ʊ][ʌ])	Strong (e.g., <u> → [u])

2.2.1.3 Insights from previous research on visual word recognition

Listeners, especially literates, can also exploit orthographic cues in speech perception and spoken word recognition. Orthographic information can also be encoded in the lexical representation along with phonological and semantic information. Also, as discussed in the previous section, one major difference between the logographic and alphabetic writing system is whether the phonological code can be accessed directly from the orthographic form. Given this difference, one question is: how do people using different writing systems access lexical representations? More specifically, is lexical access mediated by the orthographic or phonological information? Also, is there a difference between logographic and alphabetic language speakers in terms of how phonological and orthographic information is activated in lexical access? The study of visual word recognition may provide us insights for the role of orthographic and phonological information in lexical access and lexical representation, which is crucial in the learning of new words – learners of different writing systems may rely more on the orthographic or phonological code depending on the nature of learner’s L1 writing system.

Next, several studies of visual word recognition that discuss the phonological and orthographic activation during the spoken word recognition processes will be reviewed.

Perfetti and Zhang (1991) claimed that phonological information is activated only after the whole character and meaning is recognized in Chinese. Conversely, phonological information of an alphabetic word is activated prior to the activation of semantic information, as is later supported by the studies done by Lukatela and Turvey (1994) and Lukatela, Frost, and Turvey (1998).

In a series of word naming experiments conducted by Lukatela and Turvey (1994), they used three types of prime-target pairs: semantically related primes (toad – frog), homophonic to the semantically related primes (e.g., towed – frog), pseudohomophonic to the related primes (e.g., tode – frog). It was found that at short prime exposure (50 ms), all three types of primes produced equal priming to the naming of the target word (e.g., frog) for native speakers of English. However, orthographic primes, such as *told* and *tord* did not produce any priming effect, so they concluded that the processing of phonological information should mostly occur in the initial stage of lexical access during English speakers' visual word recognition. Moreover, Lukatela, Frost, and Turvey (1998) further investigated the priming of pseudohomophones, and they found a significant priming effect at 29 ms of stimulus onset asynchrony (i.e., SOA, measured from the onset of the prime until the onset of target), which was thought to be too brief for phonological activation. However, it should be noted that these experiments were conducted on English native speakers who use an alphabet. One might wonder whether Chinese speakers who use a logographic writing system would show the same pattern of results (i.e., phonological activation at the early stage of visual word recognition).

I now turn to the potential difference in the processing of orthographic information between alphabetic and non-alphabetic readers.

In Huang and Hanley (1995), children from Britain, Hong Kong and Taiwan were tested with a phonological awareness test and a visual skill test. The results indicated that, when children's IQ and vocabulary size were matched, phonological awareness was significantly related to reading ability in British children, while visual skills were crucial in the reading ability of children in Taiwan and Hong Kong. This finding further supports the hypothesis that Chinese language learners do rely more on visual information while learning to read, whereas English language learners rely more on phonological information of the words in reading.

Wang et al. (2003) investigated the role of phonological mediation in Chinese and Korean speakers' visual word recognition of English words. In their study, subjects had to decide as quickly as possible whether a stimulus word belongs to a semantic category shown earlier (e.g., *A flower* → *Rows/Rose*). The stimuli contained both similarly and dissimilarly spelled homophones to the target exemplars (e.g., *feat* vs. *feet* as a body part; *serial* vs. *cereal* as a food). Their results showed that Korean subjects made more errors in the homophone condition (e.g., *feat* for body parts) than on phonologically unrelated spelling controls (e.g., *fees*) regardless of spelling similarity. Their confusion arose from their dependence on phonology, whereas for the Chinese group there was a strong effect of spelling similarity. That is, Chinese subjects were more accurate in the less similarly spelled condition (e.g., *serial vs. cereal* as a food) regardless of phonological similarity, which implies that they were more attentive to orthographic information. This finding suggests that, compared to a logographic L1, an alphabetic L1 background has a facilitative effect on recognizing words visually in an alphabetic L2. In addition, even though Korean words are arranged in syllabic blocks, this does not affect the Koreans' reliance on phonology in visual word recognition.

Models of visual word recognition are mostly based on alphabetic writing systems, such as the Interactive Activation (IA) model proposed by McClelland and Rumelhart (1981). Based on the IA model, Dijkstra and Van Heuvan (1998, 2002) proposed the Bilingual Interactive Activation (BIA) model (see Figure 2.2 below) to account for bilingual visual word recognition.

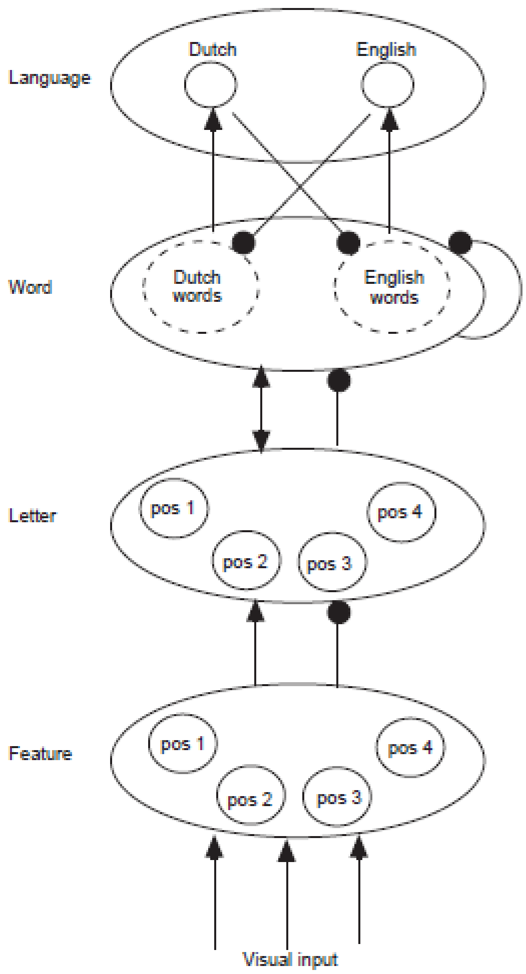


Figure 2.2. Bilingual Interactive Activation model. Reprinted from "The architecture of the bilingual word recognition system: From identification to decision " by T. Dijkstra and W. J. B. Van Heuven, 2002, *Bilingualism: Language and Cognition*, *5*(03), p. 177. Copyright 2002 by Cambridge University Press. Reprinted with permission.

The major difference between IA and BIA is that there are two lexicons in the word and language level. For example, when a bilingual speaker recognizes words visually, the Dutch lexicon will inhibit the activation of English words, and vice versa. On the other hand, Dutch words will activate the corresponding Dutch lexicon, and likewise English words will activate the corresponding English lexicon. This model is mainly built upon alphabetic writing systems, which have relatively more transparent grapheme-to-phoneme correspondence than a logographic writing system, such as Chinese as discussed earlier. The models based on alphabetic writing systems cannot be completely applied to a logographic writing system like Chinese, since Chinese does not have letter units, features and the phoneme-letter correspondence. Ehrich, Zhang, Mu, and Ehrich (2012) also pointed out that models derived from alphabetic orthography are not sufficient to account for logographic readers' visual word recognition process.

Dijkstra and Van Heuven (2002) further proposed a modified BIA model: BIA+, in which the letter and feature nodes are consolidated into two orthographic levels of processing: lexical orthography and sublexical orthography. The most notable difference between BIA and BIA+ is that BIA+ might be suited to account for recognition process differences between two distinct writing systems, such as Chinese and English.

In sum, the radical difference between a logographic and alphabetic writing system may induce different processing strategies in visual word recognition: the extent to which – and the time point at which – phonological information is activated during word recognition may depend on the L1 writing system.

Looking at this issue from a different perspective, in spoken word recognition, could orthographic information also be activated in literate listeners? Studies on visual word recognition investigate phonological activation patterns when reading a single word (i.e., from form to sound), while studies on the effect of orthography on speech perception investigate the possible orthographic activation during spoken word recognition as well as the possible interaction between orthographic form and auditory input. For example, if a listener whose L1 alphabet is transparent (e.g., Spanish) is learning a new word in a new alphabetic language, will orthographic form facilitate or inhibit learning when s/he hears a word whose orthographic form mismatches with its pronunciation (e.g., seeing <bokat> while hearing [bukat])? Also, do the users of logographic writing systems (e.g., Chinese) rely more on the orthographic information when learning an unfamiliar word in an alphabetic language (e.g., English)? Based on previous findings from visual word recognition studies, logographic users, such as Chinese, would rely more on the visual/orthographic forms to retrieve meanings than non-logographic users, so it is hypothesized that orthographic information might be beneficial for Chinese speakers to form lexical representations. On the contrary, for alphabetic language speakers (e.g., English, Spanish, Korean), phonological information will carry more weight than orthographic information when lexical representations are formed.

2.2.2 The role of orthography in speech processing

2.2.2.1 The effect of orthography on speech perception

In the previous section, the main features of alphabetic and logographic writing systems in terms of the grapheme-to-phoneme correspondence as well as the differences in processing between them have been discussed. The processing of the orthographic forms not only occurs in visual word recognition but can also influence spoken language processing.

In the typical time course of language development, perception of speech sounds occurs much earlier than the development of literacy. The onset of reading and writing could have an impact on speech perception (Burnham, 2003), and poor reading ability has been found to be associated with poor performance on speech perception tasks (Brady, Shankweiler, & Mann, 1983; Joanisse, Manis, Keating, & Seidenberg, 2000; Manis et al., 1997; Mody, Studdert-Kennedy, & Brady, 1997). It has been found that orthography can influence speech perception and spoken word recognition (Dornbusch, 2012; Jakimik, Cole, & Rudnicky, 1985; Dupoux & Mehler, 1992; Ziegler & Ferrand, 1998; Ziegler et al., 2004; Chéreau et al., 2007; Detey & Nespoulous, 2008; Taft et al., 2008; Pattamadilok et al., 2010) and also speech production (Alario et al., 2007; Damian & Bowers, 2003; Rafat, 2011). One question about the role of orthographic information in speech perception and spoken word recognition is whether orthography comes into play prelexically or postlexically (i.e., before or after the meaning of the word is accessed). If prelexically, the mediated mapping from speech signal to lexical representation will be supported. On the contrary, it is also possible that orthographic information is activated postlexically or at the same stage of lexical retrieval, in which case evidence from experiments tapping into the higher level of representation (i.e., lexical representations) will be needed. One of the studies that may support the claim that orthographic information is activated at the same stage of lexical retrieval is that of Ziegler and Ferrand (1998) in which they inverstigated whether spelling consistency would affect speech perception. In an auditory lexical decision task, half of the items consisted of French words where the rimes could be spelled in multiple ways, while the other half consisted of French words where the rimes could only be spelled in one way. Their results demonstrated that words containing the rimes with multiple possible spellings were responded to slower than those with only one possible spelling, suggesting that listeners' lexical decision was influenced by the orthographic forms of words in their memory. In other words, the orthographic forms of words were activated with the lexical representation during the process of spoken word recognition.

Dijkstra, Roelofs, and Fieuws (1995) found that in a phoneme monitoring task, when listeners were asked to detect target phonemes in spoken words or nonwords, Dutch phonemes with a secondary spelling yielded slower responses in phoneme monitoring than those with only one spelling. Although the task was to detect the phoneme, the orthographic form of the whole word could also be activated. Listeners might still use the orthographic form of the whole word as a basis for phoneme detection, so it is possible that the orthographic activation occurs at the lexical level.

However, Chéreau et al. (2007) found evidence in favor of prelexical activation of orthographic information. In a series of priming experiments, they manipulated the degree of orthographic overlap in the rimes between the targets and primes. For example, the pair "dream" and "gleam" has both phonological and strong orthographic overlap in the rime, whereas the pair "dream" and "scheme" also has both phonological and orthographic overlap in the rime except that the degree of orthographic overlap is weaker than that of "dream" and "gleam". They found that in an auditory lexical decision task the priming effect is a function of the degree of orthographic overlap – the stronger the overlap, the stronger the priming effect.

Cognitive neuropsychological studies also showed the effect of orthography on speech perception. For example, Pattamadilok et al. (2010) used the transcranial magnetic stimulation (TMS) technique with an auditory lexical decision task in which the spelling consistency was manipulated (consistent vs. inconsistent) to investigate the neural mechanisms of the orthographic effect in speech perception. They found that when the brain region responsible for orthographic processing (left ventral occipitotemporal cortex) was interfered with (when TMS was on), there was still a significant effect of spelling consistency (i.e., faster RT for consistently spelled words than inconsistent ones). However, no advantage of spelling consistency was found when phonological processing was suppressed by TMS, suggesting that orthographic effects are associated with phonological processing.

Taken together, it is clear that orthography can influence spoken word recognition and speech perception, yet it remains debatable whether the activation of orthographic information occurs pre- or postlexically. If prelexically, a mediated mapping from speech signal to lexical representation will be supported. As Gaskell (2007) noted, most connectionist and statistical models of spoken word recognition assume a prelexical level of lexical access. For example, TRACE (McClelland & Elman, 1986), Shortlist (Norris, 1994), Distributed Cohort Model (Gaskell & Marslen-Wilson, 1997) and PARSYN (Luce, Goldinger, Auer, & Vitevitch, 2000), assume a prelexical level of phonemes and features as processing units. However, the role of orthography in speech perception is overlooked in these models. As mentioned earlier, orthographic information can influence speech perception, so in the models of spoken word recognition, the processing units (e.g., phonemes, syllables) could also be associated with orthographic forms, especially for literates. In the next section, I will discuss how orthographic and auditory input may interact with each other, which provides an alternative view on spoken word recognition.

2.2.2.2 Orthographic and auditory inputs

Although it has been established that speech perception and spoken word recognition can be influenced by the orthographic forms (Petrova, Gaskell, & Ferrand, 2011; Ziegler & Ferrand, 1998; Ziegler et al., 2004; Ziegler, Muneaux, & Grainger, 2003), most models of spoken word recognition do not take the orthographic information into account, possibly because activation of the orthographic information is unnecessary for people who are illiterate or children before literacy training. However, as mentioned earlier, orthography can affect speech perception and should be acknowledged in the models of spoken word recognition and lexical access. One possible solution would be to include orthographic information as an optional layer of processing in lexical access and spoken word recognition.

On the other hand, most models of visual word recognition only emphasize visual input, such as the BIA model (Dijkstra & Van Heuven, 1998, 2002) and the IA model (McClelland & Rumelhart, 1981; Rumelhart & McClelland, 1982) as discussed earlier. Frost and Ziegler (2007) also indicated that the orthographic information starts to "append to the existing connections between spoken words and semantic meanings" (p. 107) in the mental lexicon as the development of literacy progresses. Only a few models of visual word recognition have a component of auditory input. Some researchers have also attempted to find the connection between visual and auditory word recognition. Kouider and Dupoux (2001) used a cross-modal priming paradigm to investigate whether a orthographic form can prime the corresponding spoken word under masked conditions. Their results revealed that the priming effect of the orthographic prime was observed only when the primes were consciously perceived. No priming effect was found when the prime duration was too short and did not come into conscious perception. However, the primes in this study were always visual, so it is unclear whether an auditory prime will produce similar priming effect when the primes are consciously perceived. But their results complement the finding that speech perception can be influenced by the orthographic forms.

Given the findings that orthographic forms can influence spoken word recognition, an interactive activation model that includes the interaction between auditory and visual inputs is needed to account for bimodal word recognition. In the bimodal interactive activation model (Grainger & Ferrand, 1994, 1996; Grainger & Ziegler, 2008, 2011; Diependaele, Ziegler & Grainger, 2011), both visual and auditory inputs are integrated into the model (see Figure 2.3 below).

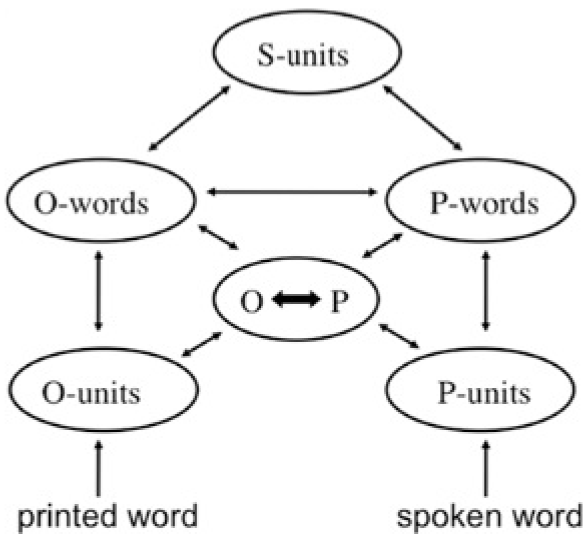


Figure 2.3. The Bimodal Interactive Activation Model (BIAM). Reprinted from "A dual-route approach to orthographic processing" by J. Grainger and J. Ziegler, 2011, *Frontiers in Psychology, 2:54*, p. 2. Copyright 2011 by Grainger and Ziegler.

In this model, there are two lexicons: phonological words and orthographic words. The interactive activation between the spoken input and the orthographic input allows the two to communicate with each other, so that orthographic representation may still be activated even without actually seeing the spelled forms. And according to Kouider and Dupoux (2001)'s findings from cross-modal priming, the interactive activation between the phonological and orthographic words should occur at the conscious level. For example, when no orthographic information is available, a spoken word [bit] in the syntactic context "*They* ___ *the enemy.*" may activate a corresponding orthographic word <beet> or <beat>, and the top-down semantic information, which could derive from the syntactic context above, will activate <beat>.

However, in some cases, the top-down semantic information may not be available (e.g., recognizing a nonword). For example, for an English reader who reads a nonword <deap>, at least two possible phonological words could be activated initially: [dip] or [dɛp], considering that the grapheme <ea> can correspond to either [i] or [ɛ], but the auditory input [dip] will inhibit the activation of [dɛp]. Based on this assumption, when the orthographic input can correspond to several possible phonological words (i.e., one grapheme to multiple phonemes, or deep orthography), the processing time will be longer than when the orthographic input only corresponds to one possible phonological word (i.e., one grapheme to one phoneme, or shallow orthography).

One drawback of this model is that it does not take the interaction between L1 and L2 orthography or the mismatched grapheme-to-phoneme correspondence into consideration. For instance, Bassetti (2008) discussed several cases where the mismatched grapheme-to-phoneme correspondence leads to non-native like production. So, if a beginning English learner of Russian hears a Russian word [bus] and sees the corresponding Russian spelling <бус>, the letter <y> corresponds to [u] in Russian but not in English; and the letter <c> corresponds to [s] in Russian but not in English (except for *–ce*). When the grapheme-to-phoneme correspondence mismatches, the visual input will inhibit the activation from the auditory input and vice versa. In other words, in this scenario, mutual inhibition, instead of interactive activation, will occur between the visual and auditory inputs. Thus, successful learning of the new grapheme-to-phoneme correspondence will depend on the suppression of the L1 grapheme-to-phoneme correspondence.

2.2.2.3 The effect of orthography on phonological acquisition and word learning

A substantial number of studies have demonstrated a link between orthography and second language phonological acquisition or word learning (Bassetti, 2008; Bassetti & Atkinson, 2015; Davidson, 2007; Detey & Nespoulous, 2008; Erdener & Burnham, 2005; Escudero, 2015; Escudero et al., 2008; Escudero, Simon, & Mulak, 2014; Hayes-Harb et al., 2010; Kaushanskaya & Marian, 2009; Rafat, 2011, 2013, 2015; Showalter, 2012; Showalter & Hayes-Harb, 2013, 2015; Silveira, 2007; Steele, 2005; Young-Scholten & Langer, 2015). What is noteworthy is that previous studies have shown mixed results as to whether orthography facilitates or interferes with learning.

In the study done by Escudero et al. (2008) using an eye-tracking paradigm, Dutch learners of English who were presented with the spelled forms of the English contrast /æ/ and /ɛ/ looked at pictures whose corresponding words contain /ɛ/ when hearing [ɛ] targets. However, those who did not see the spellings (auditory-only) looked at both the picture of word containing /æ/ and the picture of word containing /ɛ/, suggesting that exposure to L2 orthography could help L2 learners distinguish non-native phonemic contrasts.

In Hayes-Harb et al. (2010), they investigated whether presenting orthography to L2 learners can facilitate learning new words and whether mismatched grapheme-to-phoneme correspondence would influence learners' learning outcome. In a word learning experiment where participants learned novel words in an invented language, monolingual English native speakers were divided into three subgroups in the learning phase. The first subgroup saw the English orthographic form that matched the auditory words (i.e., English nonwords paired with pictures) they heard (e.g., <kamad> for the word [kɑməd] 'envelope'>. The second subgroup saw the spellings that did not match the auditory words (e.g., <kamand> for the word [kɑməd] ‘envelope’). And the third group only saw <xxxxx> (i.e., auditory only). After learning, they had to pass a criterion test (picture-auditory word matching) to make sure they had learned the associations between the pictures and words. After the criterion test, they started the testing phase. Among the mismatched picture-word pairs, they heard some auditory words that were the phonetic realizations of the mismatched spellings they saw in the learning phase (e.g., [kɑmənd] for <kamand> in the learning). In this case, the auditory word [kɑmənd] was a wrong auditory label for the picture ‘envelope’. Their results showed that the accuracy of the mismatched spelling group was the lowest among the three subgroups, suggesting that interference from L1 orthography emerges when the auditory form mismatches the spelling conventions. In other words, when the grapheme-to-phoneme correspondence in L1 mismatches with that in L2, interference from orthography will emerge in L2 word learning.

Another similar study done by Kaushanskaya and Marian (2009) also showed interference in novel word learning from L1 grapheme-to-phoneme correspondence. In a word learning experiment, a group of monolingual English speakers and a group of early English-Spanish bilinguals had to learn an artificially constructed language where the grapheme-to-phoneme correspondence mismatched with that in English and Spanish. In the learning phase, half of the items were presented to participants both auditorily and orthographically (i.e., bimodal learning), and the other half was only presented auditorily (i.e., unimodal learning). In the testing phase, participants were asked to choose the correct English translation from five alternatives for the word they heard. They were also tested on word naming (by saying the English translations). The results showed that English monolinguals’ performance was significantly better for words that they learned unimodally than bimodally. Furthermore, English-Spanish bilinguals performed significantly better than monolinguals in words that they learned bimodally. They concluded that the experience with mismatched grapheme-to-phoneme correspondence helped bilinguals in learning a new language where the grapheme-to-phoneme correspondence mismatches with their L1. These findings further support that the orthographic information of L2 may interfere with word learning if L1-L2 orthography mapping is mismatched.

Other studies showed mixed results: orthographic information was shown to facilitate or interfere with L2 word learning depending on the orthographic depth (transparency) of the L1. Erdener and Burnham (2005) investigated whether L1 orthographic transparency influences L2 word production. In their design, two L1 groups, Turkish (transparent orthography) and Australian English (opaque orthography), were asked to repeat nonwords in Spanish (transparent) and Irish (opaque) with or without orthographic information. The results showed that production accuracy (represented as phoneme errors) was overall higher when orthography was presented to participants. However, the L1 background significantly influenced the production of two types of stimuli (opaque Irish and transparent Spanish). Turkish speakers performed significantly better than Australians when repeating Spanish nonwords, whereas Australian English speakers performed significantly better than Turkish when repeating Irish nonwords. But Australian English speakers did not show significant difference in their accuracies on Spanish and Irish. The results suggest that the orthographic transparency of L1 and L2 will determine whether orthographic input can facilitate L2 word learning. Table 2.2 below summarizes the relationship between L1-L2 orthography and its effect on L2 word production as found in Erdener and Burnham (2005).

Table 2.2. L1-L2 orthography mapping and its effect on L2 word production based on Erdener and Burnham (2005)

L1 orthography	L2 orthography	Effect on L2 word production
Transparent (e.g., Turkish)	Transparent (e.g. Spanish)	Facilitation
Transparent (e.g., Turkish)	Opaque (e.g., Irish)	Interference
Opaque (e.g., English)	Transparent (e.g., Spanish)	Null
Opaque (e.g, English)	Opaque (e.g., Irish)	Small Facilitation

Simon, Chambless, and Kickhöfel Alves (2010) investigated whether presentation of orthographic information helps English learners of French learn an unfamiliar French vowel contrast /u/-/y/ with a "new" grapheme-phoneme correspondence. In a word learning experiment, subjects had to learn monosyllabic triplets that only differed in the vowel (/u/-/y/-/i/). The auditory-only group heard the words only, while the auditory-spelling group saw the spellings and heard the words in the training phase. In their first experiment, the stimuli were produced by multiple speakers, and the results showed that the two groups were not significantly different in their accuracy on the picture-word matching task. They attributed this lack of between-group difference to the difficulty caused by variability in the stimuli. In their second experiment, the stimuli were produced by a single talker, but they still failed to find a between-group difference because of the ceiling effect found in most subjects' performance. They argued that the lack of orthographic effect could be due to the opaque English orthography, which makes native speakers mistrust the orthographic forms. Or the participants in his study were able to distinguish words containing /u/ and /y/ in the specific consonantal context in their study. However, it is also possible that the lack of orthographic effect is because the items used in the training phase were the same as those used in the testing phase, although in the testing phase participants had to match the word they heard with one of the triplet pictures.

Although previous studies have shown that orthographic information may help L2 learners learn novel phonemic contrasts (when L1-L2 grapheme-to-phoneme correspondence does not mismatch) (e.g., Escudero et al., 2008), words presented only visually may not benefit L2 learners in distinguishing the non-native phonemic contrasts. As discussed in section 2.1.5, Ota et al. (2009) found that Japanese learners still tended to mistakenly accept words contrasting in /r/ and /l/ as the same word (e.g., mistakenly accept "ROCK" as a related word to "KEY"). It is unclear whether Japanese leaners can also benefit from the orthographic forms of /r/ and /l/ if the task were conducted bimodally (i.e., both visual and auditory presentation.)

The studies reviewed in this section so far all involved processing at the lexical level. Escudero and Wanrooij (2010) further investigated the locus of the effect of orthography – whether it can also be found at the sublexical level. Furthermore, they compared Spanish beginning and advanced learners' performance in two tasks: a purely auditory XAB categorization task (AUDI) and a vowel classification task where listeners chose the spelling that corresponds to the vowel they heard (ORTH). It was hypothesized that the mismatched mapping between Spanish (a transparent orthography) and Dutch (an opaque orthography) will cause interference in vowel categorization. The results of the AUDI task showed that the Spanish learners of Dutch had lowest accuracy on the contrast /a-ɑ/. However, the ORTH task showed that the orthography of the /a-ɑ/ contrast provided a durational cue to the learners, because the grapheme <aa> corresponds to /a/ and the grapheme <a> corresponds to /ɑ/. On the contrary, in the ORTH task the learners' performance in identifying the vowel /y/ was the lowest among all but highest in the AUDI task when /y/ was paired with /i/, because in the ORTH task the learners tended to choose <u> for the Dutch vowel /y/ (<u> is a grapheme corresponding to the Spanish /u/ and Dutch /ʏ/), suggesting interference from the mismatched mapping between Spanish and

Dutch. Moreover, Spanish learners were misled by the grapheme <uu> which corresponds to /y/ in Dutch, because they mistakenly considered the double <u> as a cue for lengthening.

It is also noteworthy that the advanced learners, who have had more experience with the Dutch orthography than the beginners, overall performed better than the beginner group in the ORTH task except for the pair /y-ʏ/, for which they still had difficulties in the vowel classification task.

Escudero (2015) conducted another study to investigate the effect of orthography on the discrimination of Dutch minimal pairs involving vowel contrasts that were difficult or easy to distinguish. No effect of orthography was found when the contrasts were either very easy or very difficult to distinguish, and also when the two words did not form a minimal pair. However, orthographic information was found to be beneficial in distinguishing only two contrasts that were relatively easy within the category of difficult pairs. Her finding suggests that the benefit of orthography on the discrimination of difficult minimal pairs could be dependent upon the type of contrast: orthographic information does not help in discriminating the easiest and most difficult contrasts.

In addition to research on the effect of orthography on the acquisition of L2 phonemes, Showalter and Hayes-Harb (2013) investigated whether tone marks could help L2 learners with no background in Mandarin tones learn words containing tones. Their results indicated that learners who saw tone marks while listening to the words during learning scored significantly higher in a picture-word matching task than those who did not see the tone marks. Their findings expand the previous findings that orthographic information may help L2 learners in learning segmental contrast (when L1-L2 grapheme-to-phoneme correspondence matches) to the learning of suprasegmentals. However, considering that the study conducted by Showalter and Hayes-

Harb is, to my knowledge, the first one investigating the benefit of orthography on the learning of tones, more research needs to be done on other types of prosodic cues to examine whether similar effect can be found. For example, it has been established that French listeners are insensitive to stress, which is not a contrastive feature in French (i.e., stress deafness) (Dupoux, Pallier, Sebastian, & Mehler, 1997; Dupoux, Peperkamp, & Sebastián-Gallés, 2001; Dupoux, Sebastián-Gallés, Navarrete, & Peperkamp, 2008). A possible direction would be to adopt the paradigm developed by Showalter and Hayes-Harb to investigate whether stress marks can improve French listeners' performance in discriminating stress placements.

2.3 The Current Study

The studies reviewed thus far demonstrated that L2 orthographic information may facilitate or inhibit word learning depending on the L1-L2 orthography mapping. Also, orthography has an effect on both perception and production, so the effect of orthography should be taken into account in L2 phonological acquisition. Furthermore, in PAM-L2, Best and Tyler (2007) further suggested that the mismatched grapheme-to-phoneme correspondence could be one of the reasons accounting for L2 learners' inaccurate realization of the target sound. In their example, the letter <r> corresponds to different phonemes in French ([ʁ]) and English ([ɹ]), so English learners of French may mistakenly assimilate the French [ʁ] to English [ɹ]. However, given the exposure to orthographic information, listeners may be able to form two categories faster than those who do not have exposure to the orthographic information as shown by Escudero et al. (2008).

As discussed in the previous chapter, one of the challenges in spoken word recognition is the variation in speech. One scenario of word learning is in daily conversation where learners will encounter words that they have not learned before, or words that are pronounced in different ways by different speakers, which is also one of the challenges for second-language learners. For example, if a beginning learner of English hears the word “past” produced by an accented speaker as [pɛːst], the word “pest” instead of “past” might be activated. Learners may overcome the variability in speech given sufficient training with high variability materials (e.g., Bradlow et al., 1997, 1999), such as being exposed to words spoken by multiple talkers, but there was also recent counter-evidence showing that high-variability training may inhibit learning, especially for learners with weaker perceptual abilities (Perrachione et al., 2011). So one question is how L2 learners link two pronunciation variants to one single word. For example, how does a beginning learner of English associate both [iðɚ] and [aɪðɚ] with the word “either” if s/he has only heard one of the variants? One possible resource for listeners, especially L2 learners, to recognize words even in variable contexts and learn the novel phonological contrasts would be the orthographic information.

Based on the preceding review of previous research on the effect of orthography on the learning of L2 segmental and suprasegmental contrasts, the current study seeks to expand on these previous findings by investigating whether orthographic information can help L2 learners learn words containing phonetic variants, specifically words with free variants. In general, previous studies examining the benefit of orthography on speech learning mainly investigated whether L2 learners can become better at distinguishing an unfamiliar contrast based on the corresponding orthographic forms. However, little is known about whether learners can associate two phonetic variants with the same lexical entry when the orthographic form of the word is provided during learning. In this dissertation, using a word learning paradigm modified from

Hayes-Harb et al. (2010), I will investigate whether presenting orthographic forms to L2 learners during learning can help them associate two free variants with one single lexical entry.

Chapter 3. Experiment 1: Learning of vocalic free variation

The goal of Experiment 1 is to test whether orthographic information helps L2 learners learn a free variation of two vowels in an artificial language and also helps them in general word learning. The specific research questions to be addressed in this experiment are:

1. Can orthography help L2 learners link two free variants to one single lexical entry?
2. Does the effect of orthography vary depending on learners' L1 background, specifically, alphabetic vs. non-alphabetic (logographic)?

It is hypothesized that:

1. Given the orthographic information during the learning phase, L2 learners would be able to establish a single lexical entry for words with free variants. Learners will be better at associating a picture with two versions of a word differing in one allophone if they are exposed to the spellings than those who do not the spellings while learning the words.
2. Mandarin learners will benefit more from orthography than American learners, because: (a) Chinese and the target language do not share the same orthography, which will result in less interference from L1 orthography; and (b) Chinese learners will rely more on orthographic information in lexical access than the users of alphabetic languages as found in the studies of visual word recognition.

3.1 Design and Paradigm

A word learning paradigm (Hayes-Harb et al., 2010) was adopted for the current experiment. In this paradigm, participants were to learn words in a pseudo-language by hearing words paired with pictures. For the purpose of the current experiment, a familiarization phase was added to Hayes-Harb et al.'s original paradigm, in order to allow learners to observe the vowel alternations and whether the alternations changed meanings. Thus, each participant had to complete three phases: familiarization, learning, and testing. It was expected that learners would carry over what they had observed in the familiarization phase to the learning and testing phase. Each phase is described in detail in the following sections.

To answer the first research question, three within-subject conditions were created: *free variants* (words that have two phonetic realizations; that is, pictures that have two associated phonetic variants), *minimal pairs* (words that differ in only one sound in the same position), and *baseline* (words that differ in many ways do not involve any sound alternations). The free variant condition allows us to assess whether participants would be able to link the two free variants to one lexical entry: if they do, then they should accept the variant as a correct name for the picture during the test phase. The minimal pair condition is included to check whether participants are paying attention to the actual phonetic form (i.e., sound alternations) of the items. If they do not pay attention to the sound alternations, both a high false alarm and high hit rate would be observed – they would just consider any two similar sounding items as the same. The baseline items are used to check the general word learning outcome by all participants. Finally, in order to compare the effect of orthography on the learning of free variation, the most important experimental manipulation for the current research question is that the orthographic forms of the words are only available to one group of participants.

To answer the second research question, users of an alphabetic language, English, and users of a logographic language, Mandarin, were chosen for this study. The rationale for choosing these two groups is twofold. First, both languages are matched in terms of orthographic depth – both have opaque or deep orthography, although relatively speaking Mandarin is even more opaque than English. Second, the availability of participants of these two language groups is higher than the others on the Indiana University campus.

In summary, the current experiment employs a 2 X 3 X 2 design, which are 2 orthographic conditions: Orth+ and Orth– (where Orth+ sees the spellings and Orth– does not see the spellings during familiarization and learning), 3 Item Types (free variant, minimal pairs and baseline) and 2 L1 groups (English vs. Mandarin). Details of the design of each condition are described in the following sections.

3.1.1 Familiarization

At the beginning of the experiment, participants were informed that they were going to hear some new words in a new language, and the pictures shown on the center of the screen represented the meaning of the words they heard. One crucial manipulation was the between-subject variable: Orth+ vs. Orth–. The Orth+ group only saw the written forms of the words under the pictures and the Orth– only saw a series of Xs while hearing the words. Following the design of other studies, none of the participants were informed of the presentation of the written forms (see Figure 3.1 for sample items).

The task was self-paced: to ensure that they paid attention to the task, participants were instructed to repeat aloud each word they heard at their own pace. After they repeated the word, they were instructed to press a key on the response box to proceed to the next trial. In addition, they were told that some words may have more than one possible pronunciation. The purpose of this phase was to familiarize them with the novel words and let them observe the pronunciation rule or regularities. No feedback was provided. They were not forced to remember all words and pictures in this phase. Two variants of the same word or two words from one minimal pair did not occur consecutively.

Details on stimuli used in this phase are presented in section 3.3.5 (number of items in each phase) below.

3.1.2 Learning

The procedure in this phase was the same as that in the familiarization, except that they learned another new set of words, for which they were tested on the variant (see Table 3.1 below: they learn variant A of free-variant items and word A of the minimal-pair items, and were later tested on variant B and word B respectively in the testing phase).

Table 3.1. Sample Free Variant and Minimal Pair items presented in the learning and testing phase

Condition	Learning	Testing
Free Variant	variant A [fɔsat] ‘shark’	variant B [fusat] ‘shark’ (matched)
Minimal Pair	word A [gekaf] ‘book’	word B [gakaf] ‘book’ (mismatched)

Participants were told that they would be tested on their memory of these words afterwards, and that they should try to memorize the items. In this phase, again, no feedback was provided. After completing this phase, participants performed a criterion test in which they were tested on the association between the words and the pictures they learned in the learning phase. In each trial, they saw a picture (with no written forms) and heard a word. They needed to decide if the picture matched the word they heard.

The items for the criterion test were the same as those used in the learning phase, but half of the items were paired with correct pictures (matched condition), whereas the other half was paired with clearly wrong pictures (mismatched condition). They could not advance to the next phase (i.e., testing) until they reached 90% accuracy (adopted from Hayes-Harb et al., 2010). In other words, if they did not score at least 90% correct, the program would loop back to the learning phase and they had to learn the words again. The results showed that only 6 participants (out of forty) needed more than one cycle of learning to move on to the next phase, and no participants needed more than two learning cycles. An analysis of variance (ANOVA) showed that those who needed 2 cycles were not significantly different from those who only needed 1 cycle in both RT and accuracy on the picture-word matching task (both p values > .1).

The criterion test was implemented to make sure that participants learned the association between the words and pictures. If they were unable to correctly associate the pictures with words, it would be hard to assess their learning of free variation in the following testing phase because their low accuracy in the testing phase could be due to their poor association between words and pictures in the first place. For example, if one cannot associate the word [fusat] with the picture "shark" in the learning phase, and then if s/he responded "*no*" to the free variant [fɔsat] in the testing phase, then we cannot conclude that s/he did not learn the free variation. Instead, it can be that s/he had not learned the association between [fusat] and "shark" at the first place.

3.1.3 Testing (picture-word matching)

In this phase, participants were presented with variant B of the free variant and word B of the minimal pairs of what they learned in the learning phase (see Table 1). For example, a participant learned [fɔsat] as the auditory label for the picture "shark" in the learning phase. In the testing phase, s/he hears [fusat] and has to decide if that is the word for the picture "shark" by pressing the green key for "*yes*" and the red key for "*no*" on the response box (the expected answer for this example should be "*yes*", because [fɔsat] and [fusat] are free variants.) Their response time and accuracy were collected. The complete paradigm is illustrated in Figure 3.1 below.

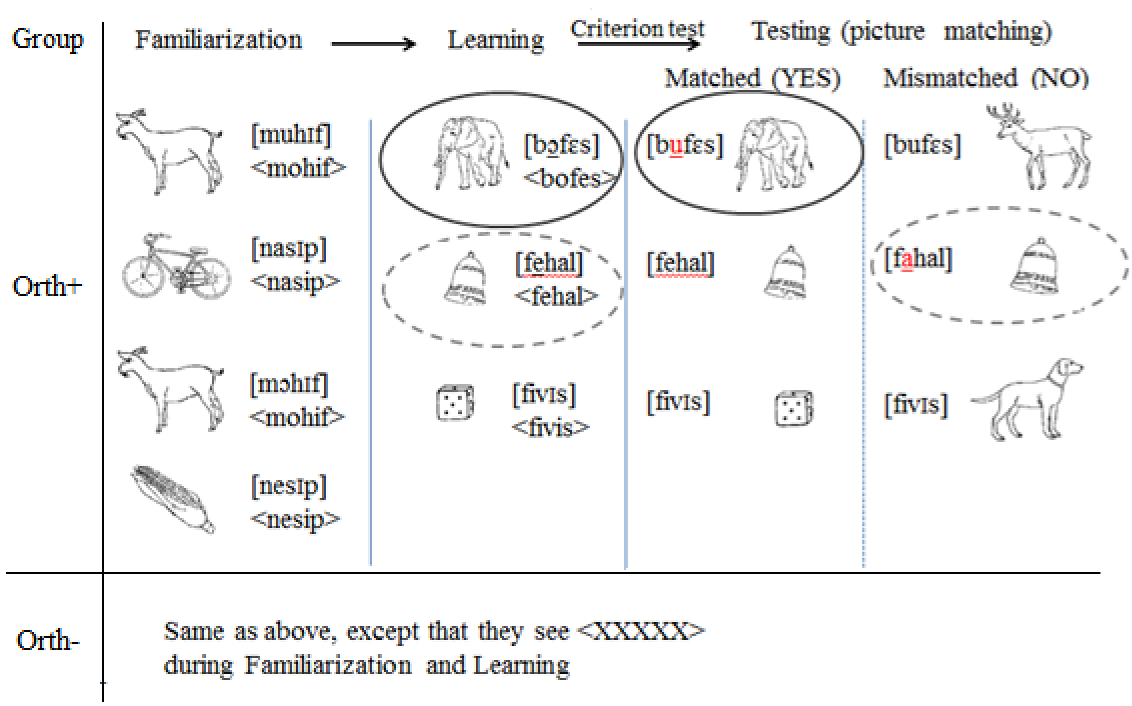


Figure 3.1. The word learning paradigm and the flow of the three phases: familiarization, learning and testing. In each phase, each picture was paired with an auditory word (shown in "[]") and spellings (shown in "< >") were presented in Orth+.

Note: The key conditions of interests are circled in the figure above: The solid circles indicate the same words, i.e., free variants, while the dotted circles indicate different words, i.e., minimal pairs.

3.2 Participants

Twenty-two Mandarin native speakers from Taiwan (Taiwanese hereafter; mean age = 29.9, *SD* = 4.98; length of residence (LoR) in English-speaking countries < 3.5 years) and 23 native speakers of American English (mean age = 20.31, *SD* = 1.36) were recruited on the Bloomington campus of Indiana University. They were paid $10 for their participation. All participants reported normal or corrected-to-normal vision and hearing. However, the data of three Americans and two Taiwanese participants were not included in the analysis: One American participant was excluded because she failed to pass the hearing screening and needed 3 learning cycles before moving onto the testing phase. Another American participant was excluded due to his background in linguistics, which could potentially confound the result because of his metalinguistic knowledge. One Taiwanese participant was excluded because her data showed a completely opposite pattern from other Taiwanese participants and that she needed 3 learning cycles to continue. Lastly, one American and one Taiwanese participant were excluded due to the unclear instructions that they received during familiarization. Hence, in the end the data of 20 Taiwanese participants and 20 Americans were included in the analysis.

Considering that mainland Chinese learns a Romanized spelling of Chinese (Pinyin) in early childhood and may have more exposure to alphabetic writing systems (Read, Yun-Fei, Hong-Yin, & Bao-Qing, 1986), Mandarin speakers from mainland China were also excluded from this study. Similarly, linguistics majors and minors were excluded from this study to minimize the possible confounding effect of linguistic knowledge.

Regarding their education level, all of the twenty American participants were undergraduate students, while only two of the Taiwanese participants were not graduate students.

Participants from each L1 group were further assigned into two subgroups: Orth+ and Orth– where Orth+ saw the spellings of words during familiarization and learning phase, while Orth– only saw a series of "X". There were ten Orth+ and ten Orth– in each language group.

3.3 Materials

Materials for this experiment were of two kinds. First, 54 novel, word-like nonwords were created (see Appendix I.) and audio-recorded for presentation in the picture matching experiment. Second, 40 black-and-white line drawings were selected from a picture database (Alario & Ferrand, 1999) (see Appendix III.) for association with the nonwords. Along with the nonwords, we decided to use pictures of familiar[1] objects (shark, book, tree, shoe, etc.), because we wanted to approximate the L2 acquisition process which consists of associating novel word forms to known objects/concepts. Therefore, rather than using both novel word forms and novel pictures (e.g., Hayes-Harb et al., 2010 or Escudero et al., 2008), we told participants that they would be learning words in a new, unknown language.

There were various constraints to stimuli construction, which will be described in turn (sections 3.3.1 – 3.3.4). In addition, there were experimental considerations regarding the number of items, which will be described in 3.3.5.

[1] Familiarity with pictures was not explicitly tested, but participants were asked by the experimenter whether they knew all of the pictures after the major experiment. If participants were unsure of a picture, they were asked what object they thought it was in the picture. Their responses were then noted by the experimenter.

3.3.1 Orthographic depth

Given that the main purpose of this experiment is to evaluate whether orthographic information helps learners associate two free variants of the same word with the same picture, it was important to ensure that the grapheme-to-phoneme correspondence was consistent. Table 3.2 below summarizes the grapheme-to-phoneme correspondence used in the current experiment.

Table 3.2. Grapheme-to-phoneme correspondence in the current experiment

Sounds	Letters
Consonants	One-to-one correspondence (e.g., [k] only corresponds to <k> but not <c>)
[ɪ], [ɛ]	<i>, <e> (in unstressed syllables)
[i], [e]	<i>, <e> (in stressed syllables)
[a]	<a>
[ɔ], [u]	<o>

In order to make the invented language a transparent one and avoid possible interference from one's opaque L1, a native speaker from a different language that also uses transparent orthography was selected to record the stimuli (see section 3.3.2 below for details).

3.3.2 Stimulus recording and talker variability

In a pilot AX discrimination task, stimuli produced by multiple talkers appeared too difficult for the participants to distinguish. When the stimuli within a trial were produced by two talkers, Mandarin participants only scored 36% (*SD* = 0.48) on "different" trials. Their accuracy improved to 61% (*SD* = 0.49) when the stimuli within a trial were produced by one single talker. Thus, in the current experiment, one single talker, who was a phonetically-trained near-native female German speaker, was recruited to record the stimuli. Stimuli were recorded in a quiet room using Marantz PMD 620 digital portable recorder. The amplitude of all items was normalized to -18.5 dB total RMS.

3.3.3 Nonwords for participants with English or Mandarin as native language

It was important to create nonwords which are not too close to real English words in order not to unfairly advantage English participants, and not too far from possible words as to remain learnable.

All stimuli were disyllabic nonwords (CV.CVC) based on the ARC nonword database (Rastle, Harrington, & Coltheart, 2002) and MRC Psycholinguistic database (Coltheart, 1981). None of the stimuli were real words in participants' native languages (English or Mandarin). All items were stressed on the first syllable. Considering the possible word-final neutralization of the voicing contrast, the word-final consonant was always voiceless.

In addition, according to the Neighborhood Activation Model (Luce & Pisoni, 1998), word naming will be slower if the word is in dense neighborhood due to the competition from other word candidates. Therefore, to control for the possible effect of neighborhood density, the English phonological and orthographic neighborhood density of these nonwords were obtained from the Speech and Hearing Lab Neighborhood Database of Washington University in St. Louis (Sommers, 2003), which is based on the Hoosier Mental Lexicon (Nusbaum, Pisoni, & Davis, 1984). No items had more than 1 phonological and 3 orthographic neighbors in the English lexicon. The average number of the phonological neighbors was 0.04 (*SD* = 0.2), and that of orthographic neighbors was 0.75 (*SD* = 1.03).

3.3.4 Choosing the vowel alternations

Inspired by the Polish *o*-raising rule (e.g., Buckley, 2001) where /o/ is realized as [u] when it precedes a word-final voiced consonant, a vowel alternation between [o] and [u] would be a typologically appropriate alternation for our purposes. The major difference between the [o]-[u] alternation selected in the current study and the Polish *o*-raising is that the alternation in the current study was context-free.

However, a pilot AX discrimination task showed that Taiwanese participants had difficulties in distinguishing the unstressed tense [o] from unstressed tense [u] produced by German speakers ($N = 12$, $M = 0.73$, $SD = 0.44$). Only one participant scored above 80% correct, and one participant scored below 20%. Also, in "different" trials (i.e., when A and X were different), Mandarin participants only scored 50% ($SD = 0.5$), suggesting that they were guessing whether the two words were the same.

For the present experiment, it is crucial that participants could at least reach 90% accuracy in distinguishing the target vowels [o] and [u]. If they failed to clearly distinguish [o] and [u], then their responses to accept either variant as a correct label for the picture in the testing phase (i.e., a picture-word matching task) could not reflect whether they really had learned that [o] and [u] were not contrastive — it might be simply because that they merged these two vowels as one single category in their perception in the first place. Thus, a less confusable pair was needed in order to see the learning effect of free variation: mapping of the two variants onto one lexical entry was not due to the fact that the variants were perceived as the same token. Thus, the pair /ɔ/-/u/ was chosen for the current experiment. In addition, all alternations occurred on the first and stressed syllable to avoid possible confounds from the difference in the perception of stressed versus unstressed syllables.

The free-variant items were words involving free variants alternating between /ɔ/ and /u/. Both words were paired with the same picture (e.g., [fɔsat/fusat] 'shark'). The minimal-pair items were minimal pairs involving the alternation of /e/ and /a/. Each word in the pair was paired with one picture (e.g., [gekaf] 'book'; [gakaf] 'something else that the participants did not learn'.)

3.3.5 Number of items in each phase

In the familiarization phase, there were 4 pairs of free variants (i.e., 4 pictures), 4 minimal pairs (i.e., 8 pictures) and 6 baseline items (i.e., 6 pictures). Each picture was repeated four times, yielding a total of 72 trials.

In the learning phase, there were in total 24 items (see Table 3.3 below for illustration). Each item was repeated four times, which gave a total of 96 trials. There were also two pairs of free variants that were presented as both Variant A and B as well as two minimal pairs that were presented as both Word A and B. The purpose of including these items was to let participants observe again the alternations so as to reinforce their learning of the free variation. Since both forms of the free variants and minimal pairs were presented, these items were not included in the subsequent criterion test and the testing phases.

Table 3.3. Number of items in the learning phase

	To be tested in the Criterion Test and the Testing phase	**Used to reinforce learning; To be excluded in the Testing phase**	**Grand total**
Free variants	4 (only Variant A)	2 pairs (both Variant A and B) = 4 items	8
Minimal pairs	4 (only Word A)	2 pairs (both Word A and B) = 4 items	8
Baseline	8	N/A	8
Subtotal	16	8	24

The criterion test consisted of items they learned in the learning phase, except for the learning-reinforcement free-variant and minimal-pair items that were excluded in the testing phase (see Table 3.4 below). Each item was paired with the correct picture once (Matched) (e.g., [fusat] 'shark'), and once with an incorrect picture (Mismatched) (e.g., [fusat] 'spiderweb'), yielding a total of 32 trials.

Table 3.4. Number of items in the criterion test

	Matched	**Mismatched**
Free Variants	4	4
Minimal Pairs	4	4
Baseline	8	8
Total	16	16

In the testing phase (a picture-auditory word matching task), participants were presented with one auditory form at a time, paired with a picture, and had to decide whether the pairing was correct (matched) or not (mismatched).

Crucially, the items learned in the learning phase were presented in their variant (or minimal pair) form in the testing phase. For example, a free-variant item from the learning phase (e.g., variant A, [bufɛs]) would appear as [bɔfɛs] (variant B) in the testing phase, paired with the same picture used in the learning phase, which creates a matched pairing. A minimal-pair item from the learning phase (e.g. word A, [gekaf]) would appear as [gakaf] (word B) in the testing phase, paired with a *same* picture used in the learning phase, which creates a mismatched pairing. All items also appeared as well as in their originally learned form (in Block 2 shown in Table 3.5 below) (i.e., variant A and word A).

In order to prevent participants from using a strategy of answering "*yes*" to globally similar sounding items, the free-variant and minimal-pair items to be tested (variant B) were always placed in the first block of trials, whereas the variants they have learned (variant A) were always placed in the second block of trials. For example, if they learned [fɔsat] in the learning phase, they would be first tested on the variant [fusat] in the first block of the testing phase. However, the baseline items, which were always old tokens, appeared in both blocks to balance the matched and mismatched trials. Table 3.5 presents an overview of number of trials in each item and pairing condition.

Table 3.5. Number of trials in the testing phase (picture- auditory word matching task)

		Block 1		Block 1 Total	Block 2		Block 2 Total
	Pairing condition:	Matched	Mismatched		Matched	Mismatched	
Item condition	Free Variant	8		8	8	8	16
	Minimal Pair		8	8	8	8	16
	Baseline	12	12	24	4	4	8
	Grand Total	20	20	40	20	20	40

Note: The highlighted cells indicate the key conditions

Notably, in block 1, all free-variant items appeared in the matched condition, because the free-variant items were associated with the matched picture of the variant they learned in the learning phase. For example, if they learned [fɔsat] for 'shark' in the learning phase, and were tested on [fusat] for 'shark' in the testing phase, the answer should be "YES"/matched.

On the contrary, all minimal-pair items in block 1 were all in the mismatched condition, because the minimal-pair items (e.g., [gekaf] vs. [gakaf]) were associated with two different pictures. Only one member of such pairs (word A in Table 3.1) was learned during the learning phase (e.g., [gekaf] for 'book'). However, since the picture for the other member of the pair (word B in Table 1: [gakaf] for something other than 'book') was not learned in the learning phase, the picture of word B in the minimal pair was not presented in the testing phase. Instead, in the testing phase, word B of the minimal-pair item (i.e., [gakaf]) was still paired with the picture of word A they learned in the learning phase to form a mismatched condition. For example, they learned word A [gekaf] for ‘book’ in the learning phase, and were tested on word B [gakaf] for ‘book’ in the testing phase, for which the answer should be “NO” (mismatched).

Both the free-variant and minimal-pair items were repeated twice, but half of the baseline items appeared twice in the mismatched condition, while the other half appeared twice in the matched condition to balance the number of items in each condition.

In block 2, both variant A and B of the free variants and word A and B of the minimal-pair items were paired with a clearly wrong picture (i.e., picture-word pairings that were completely unrelated to the vocalic alternations) to form a mismatched condition. The original forms (variant A and word A in both conditions) learned in the learning phase were paired with the correct picture to form a matched condition as well as an incorrect picture to form a mismatched condition. In addition, half of the baseline items were in the matched condition while the other half were in the mismatched condition. In order to balance the number of items in the matched and mismatched, the original forms of free variants and minimal pair in matched condition appeared twice.

3.4 Procedure

All procedures described in this dissertation were approved by the Indiana University Institutional Review Board. Participants were seated in a quiet lab. As soon as they arrived in the lab, after agreeing to participate, they were briefed by the experimenter about the purpose of this study and what they would need to do. Then they first began with an AX discrimination task in order to test whether they could discriminate the target free variants [ɔ] and [u]. The items used in the AX task were the same as the test items (i.e., free variants and minimal pairs) used in the word learning experiment (see Appendix I.). The results showed that the mean accuracy on the target pair is 99% (*SD* = 0.11). After the AX task, a pure tone hearing screening test (Reilly, Troiani, Grossman, & Wingfield, 2007) was administered to ensure that participants had normal hearing. However, note that the hearing screening was not taken as a “strict” measure. In other words, participants were not automatically screened out of the study if they failed the hearing screening. Given the ambient noise and hardware limitations in the lab, it was likely that the hearing screening was not as accurate as it was designed to be. Therefore, if a participant failed the hearing screening and his/her performance in the experiment showed a clearly different pattern from others’ results, his/her data would then be dropped from the analysis. But if someone failed the hearing screening and showed similar performance to other participants, his/her data would be kept.

After the hearing screening, they filled out a language background questionnaire. After filling out the questionnaire, they started the actual word learning task. The AX task was administered at the beginning of the session and then followed by the questionnaire to minimize possible priming effect from the items in the AX task in the learning experiment. All experiments were programmed in E-Prime 2.0 (Schneider, Eschman, & Zuccolotto, 2002). Participants wore a Sennheiser PC350 passive noise-attenuation headset and sat in front of a LCD monitor. They were instructed to use the E-prime response box to make their responses.

After participants completed all the experiments, they were given a printed list of the stimuli used in the word learning experiment and asked to rate how likely the item could be a real word in any language they know on a Likert scale from 1 (unlike a word) to 5 (very close to a word). The mean rating for the items in the learning phase by American participants was 2.93

(*SD* = 1.33), and 3.22 (*SD* = 1.4) by Taiwanese participants. For those in the testing phase, the mean ratings by American participants was also 2.93 (*SD* = 1.3), and 3.21 (*SD* = 1.34) by Taiwanese participants. A mixed-effects model was fitted to examine whether the ratings between the two L1 groups (i.e., Taiwanese vs. American) were significantly different. The results are shown in section 3.6.2 below.

3.5 Predictions

If participants learn the free variation according to which the [ɔ-u] alternation does not cause a change in meaning, then they should be able to accept the free variant B [fusat] as a correct match for 'shark' (hit) and reject word B [gekaf] when paired with the picture of 'book' (correct rejection). The accuracy on these two particular conditions (i.e., free-variant items in the matched condition and minimal-pair items in the mismatched condition; highlighted in Table 3.5 above) is of key interest: If the accuracy is low on the minimal-pair items but high on the free-variant items, then one cannot conclude that they have learned the free variation, because such a pattern would suggest that the person responded "yes" to most trials. This outcome would result in high hit rate for the matched (free variant) items and at the same time a high false alarm rate for the mismatched items minimal pair), leading to a very low *d'*. In other words, it is possible that such a pattern reflects a very liberal strategy to say "yes" without paying attention to the exact form of the stimuli (i.e., both [ɔ-u] and /e-a/ vowel differences – or possibly any vowel differences – are ignored). By contrast, if the accuracy on the minimal pair item is near chance level (50%)[2] and high on the free variant, then we might conclude that they have learned the free

[2] Note that this is a relatively difficult task, as shown in the results later, subjects who start guessing the answers to the minimal-pair items seem to know that the minimal-pair items are different from the free-variant items in some way.

variation "to some extent"– they are not just answering "yes" (matched) to all similar sounding items. Instead, this pattern would indicate that participants noticed that the alternating vowels in the minimal pair items cause meaning change while the alternating vowels in the free-variant items do not.

3.6 Results

3.6.1 Data structure and preparation

Participants' responses in the first block of the testing phase, which is the crucial experimental condition in the current study, were analyzed. Both accuracy and response time (RT thereafter) data of 40 trials per subject (i.e., 1600 observations in total) were collected and analyzed. Forty (out of 45) participants' data were analyzed. (See section 3.2 for detailed explanation for the criteria of screening out the participants.) Before fitting the model for the RT data, a histogram of the data showed that the data distribution was skewed. This was in violation of normal distribution, so the RT data was log-transformed in order to fit the model. The log-transformed RT data was close to normal distribution.

The independent variables include: L1 groups (Group hereafter), Item Type (i.e., free variants, minimal pairs, and baseline items), pairing conditions (i.e., matched vs. mismatched; Condition hereafter), and orthographic conditions (i.e., Orth+ vs. Orth–; OrthoCond hereafter). Before analyzing the data, in order to trim extremely slow responses, a univariate analysis of variance (ANOVA) with log-RT as the dependent variable and subjects and Item Type (free variant, minimal pair and baseline) as fixed factors was run to obtain the standardized residuals (*Z* residuals) of the log-RT data. The trials with *Z* values above or below 3 were then trimmed. The rationale of trimming by residuals from subjects and Item Type is that it accounted for the fact that the difficulty level of each condition might vary across subjects considering individual differences in learning ability. In the end 19 trials of all data (1.1%) were trimmed.

In section 3.6.2 as well as 3.6.3 below, I will first present the observed means and standard deviations in tables followed by bar graphs where the means are estimated marginal means from the statistical models.

3.6.2 Accuracy data

The accuracy data is coded 1 for correct responses and 0 for incorrect responses. Given this binary categorical data structure, a generalized estimated equation model was fitted to examine the effect on accuracy of the factors Group, OrthoCond, Condition, and Item Type.

It is possible that words with a higher wordlikeness rating could be learned more easily and be responded to more accurately and faster. Responses for the different items were also examined for consistency. The wordlikeness ratings were used to examine whether words with higher ratings yielded higher accuracies and faster responses. Before analyzing the effects of the fixed factors on accuracy, the wordlikeness ratings were examined separately for the learning and the testing phase. For the learning phase, a linear mixed-effects model declaring Group and Item Type as the fixed factors, Subject and Item as the random factors and wordlikeness ratings as dependent variable was fitted to examine whether wordlikeness ratings for the items in the learning phase were comparable between Taiwanese and American participants as well as between different item types. Another linear model was fitted to examine the effects of the same factors on the ratings for the items in the testing phase. The result showed a main effect of Item Type on the wordlikeness rating in both the learning phase [$F(2, 14) = 14.43$; $p = .001$] and testing phase [$F(2, 13) = 12.23$; $p = .001$]. No main effect of Group on wordlikeness ratings for items in both the learning and testing phase was found [Learning phase items: $F(1, 39) = 1.55$; $p = .220$. Testing phase items: $F(1, 39) = 1.68$; $p = .202$]. Given the significant correlation between Item Type and the wordlikeness rating, the wordlikeness rating was not included in the model for the accuracy data considering the multicolinearity between the two factors. In other words, if Item Type is found to be significant on accuracy, it could be that it is the wordlikeness ratings that cause the effect due to the correlation between Item Type and the ratings.

Therefore, in order to examine whether accuracy on different item types was an artifact of wordlikeness ratings, two generalized estimated equation models, one of which included LearningRating (wordlikeness ratings for the items in the learning phase) and another included TestRating (ratings for the items in the testing phase) with the fixed factors in this study (Group, OrthoCond, and Item Type) were fitted to see whether the effect of Item Type remained significant when either LearningRating or TestRating was held constant in the model for accuracy. The results showed that Item Type was still a significant factor on accuracy even when LearningRating or TestRating was held constant [$\chi^2(1) = 51.58$; $p < .001$ and $\chi^2(1) = 29.2$; $p < .001$, respectively]. Given that Item Type was still a significant factor even when wordlikeness ratings were held constant, wordlikeness ratings (LearningRating and TestRating) were not included in the final models.

However, there were two items rated above 4 (out of 5) (<panek> rated at 4.75 on average by Americans and <fivis> rated at 4.13 by Taiwanese). A comparison between the fitted models of accuracy including and excluding these two items showed no difference in significance levels, but for the fitted model of response times, a slight change of significance level of the three-way interaction (OrthoCond by Group by Item Type) was found: *p* value changed from .02 (when those two items were included) to .059 (when those two items were excluded). Although this three-way interaction is only marginally significant after excluding those two items, none of the significance levels in the pairwise comparisons changed. Thus, in order to keep as much data as possible given that the sample size was not very large, <panek> and <fivis> (both of which are baseline items) were still kept in the final analyses.

In addition, a post-hoc item analysis revealed that one of the free-variant items (mohil/muhil) in the testing phase and an item (differing in the coda consonant) in the familiarization phase formed a minimal pair by accident (mohif/muhif). An ANOVA showed that this item caused significant faster responses ($p < .05$) by American participants (but not Taiwanese) than other free-variant items in the testing phase. To account for this effect, this item (muhil) was coded as "1" and other items were coded as "0" as a new variable "FamMinimalPair" (i.e., whether the item had a minimal pair counterpart in the familiarization), which was then included in the mixed-effects model for the RT data and the generalized estimated equation models for the accuracy data.

Table 3.6 below presents the mean accuracy in the first block (which contains the key experimental conditions) for each L1 group and orthography condition for the test items.

Table 3.6. Mean accuracy on the test items (minimal pairs and free variants)

Test items		Orth–		Orth+	
L1 Group	Condition	Mean	SD	Mean	SD
American	Minimal Pairs	0.13	0.33	0.15	0.36
	Free Variants	0.84	0.37	0.91	0.28
Taiwanese	Minimal Pairs	0.41	0.50	0.39	0.49
	Free Variants	0.80	0.40	0.84	0.37

Table 3.7 below shows the mean accuracy on the baseline items in both matched and mismatched conditions by both American and Taiwanese participants, and in both orthographic conditions.

Table 3.7. Mean accuracy on the baseline items

Baseline items		Orth–		Orth+	
L1 Group	Condition	Mean	SD	Mean	SD
American	Matched	0.97	0.18	0.99	0.09
	Mismatched	0.97	0.18	0.97	0.18
Taiwanese	Matched	0.97	0.18	0.96	0.20
	Mismatched	0.96	0.20	0.97	0.16

A generalized estimated equation model declaring the dependent variable accuracy, fixed factors Group (Taiwanese, American), OrthoCond (Orth+, Orth−), Item Type (free variants, minimal pairs and baseline items), Condition (matched vs. mismatched) and FamMinimalPair (to account for an accidental minimal pair found in the familiarization and learning phase) and random factors (subjects and items) was fitted to examine the main effects and possible interactions of the fixed factors on accuracy rates. The results showed that the factors Group [$\chi^2(1) = .354$; $p = .552$], OrthoCond (i.e., Orth+ and Orth–) [$\chi^2(1) = 1.881$; $p = .170$] and Condition (i.e., matched vs. mismatched) [$\chi^2(1) = .163$, $p = .687$] were not significant, but Item Type (i.e., Free Variants, Minimal Pairs and Baseline) was significant [$\chi^2(2) = 183.287$; $p < .001$] in which the baseline items yielded the highest accuracy (*M*[3] = .98), the free variants yielded the second highest (*M* = .86), and the minimal pairs yielded the lowest (*M* = .33). No significant

[3] The means reported in the inferential statistics are based on the estimated marginal means from the linear mixed-effects model, while the means reported in the tables are the observed means.

interactions among any of the fixed factors were found. Figure 3.2 below shows the mean proportion correct (%) on the free variants and minimal pairs for both L1 groups, as a function of orthographic exposure.

A post-hoc pairwise comparison between the two L1 groups on the minimal pairs using Bonferroni correction showed that Taiwanese scored significantly higher than American participants on minimal pairs in both Orth– (p = .001) and Orth+ (p = .006), as shown in Figure 3.2 below. However, no benefit of orthographic information was found: Orth+ did not outperform Orth– in either L1 groups on both minimal pairs and free variants, although Orth+ of both L1 groups scored slightly higher than Orth– on the free variants.

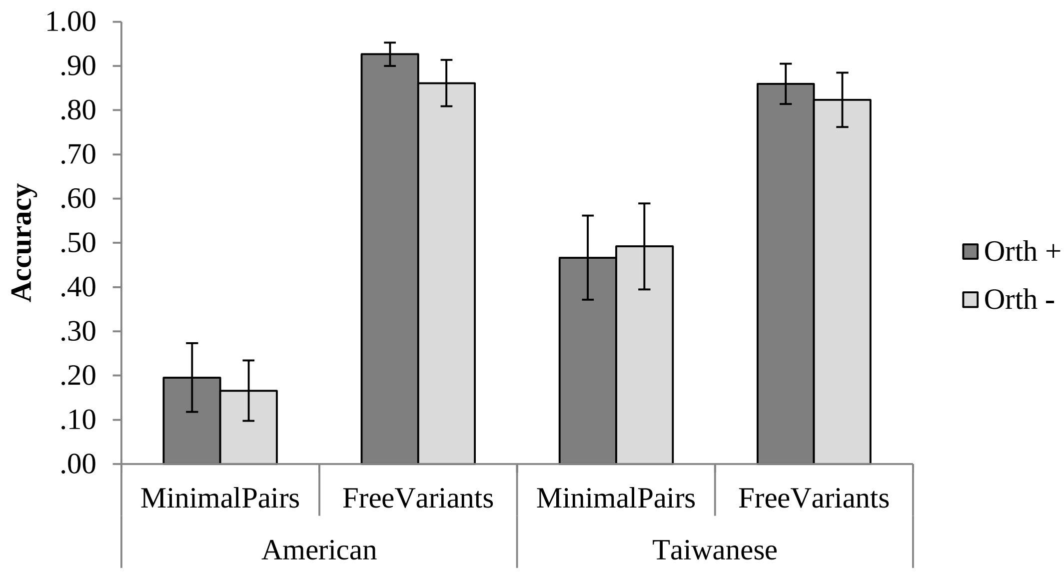


Figure 3.2. Mean proportion correct on the free variants and minimal pairs for both L1 groups as a function of orthographic exposure (dark bars: Orth+; light bars: Orth–)[4]

[4] The bar graphs are based on the estimated marginal means and standard errors from the generalized estimated equation models for accuracy and mixed effect models for RT.

In addition, considering that the letter-to-sound correspondence was mismatched for half of the free variant items (e.g., half of the items contained the target vowel [u], which was spelled as <o>), another generalized estimated equation model was fitted to examine whether there was a significant difference between the free variant [u] spelled as <o> and [ɔ] spelled as <o>. The result showed that there was no significant difference between these two type of free-variant items [$\chi^2(1) = 0.79$; $p = .778$] on accuracy.

Looking at Figure 3.2 above, one might ask: how much did learners actually learn? The learning of free variation could be reflected by the difference in accuracy between the minimal pairs and free variants: The smaller the difference between the two conditions *and* the higher the accuracy on the minimal pairs, the greater the degree of learning. Although Taiwanese participants outperformed Americans on the minimal pairs and we can conclude that Taiwanese learned the free variation better than Americans, Taiwanese also did not score above chance on the minimal pairs. Possible accounts for American participants' poor performance on the minimal pairs will be discussed in section 3.7.

Figure 3.3 below shows the mean proportion correct (%) on the baseline items for both L1 groups as a function of condition.

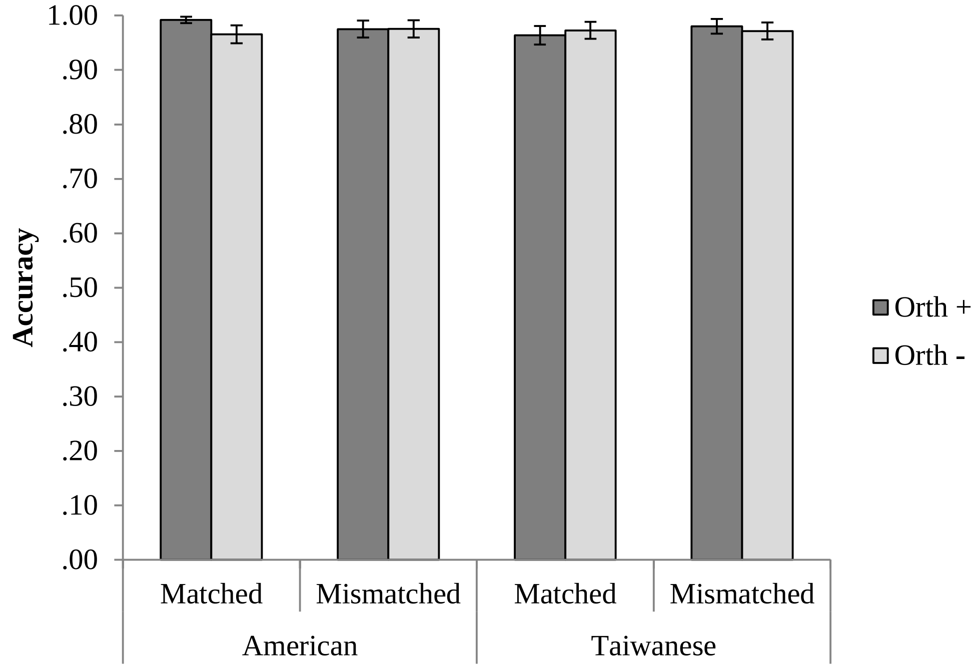


Figure 3.3. Mean proportion correct on the baseline items for both L1 groups on the matched and mismatched condition, as a function of orthographic exposure (dark bars: Orth+; light bars: Orth–)

From Figure 3.3 above, it can be observed that both L1 groups' accuracy on the baseline items was very high, and that both orthographic groups performed equally well regardless of the presentation of the written forms of the words. Post-hoc pairwise comparisons using Bonferroni correction from the generalized estimated equation model reported earlier showed that there was no significant difference between the two orthographic conditions across both L1 groups and pairing conditions ($p > .05$)

To examine participants' speed in making the responses, in the next section the RT data and results will be reported.

3.6.3 RT data

The RTs of all responses were included in the analysis. As mentioned earlier in the accuracy data, a main effect of Item Type on the wordlikeness rating in both the learning phase and testing phase was found. Thus, another two linear mixed-effects models were fitted to examine whether Item Type would remain significant when either LearningRating or TestRating was held constant in the model for RT data. The results showed that Item Type was still a significant factor on log-RT even when LearningRating or TestRating were included in the model [$F(2, 54) = 15.85$; $p < .001$ and $F(2, 64) = 15.22$; $p < .001$, respectively]. Given that Item Type was still a significant factor even when wordlikeness ratings were held constant, wordlikeness ratings (LearningRating and TestRating) were not included in the final models.

Tables 3.8 (for test items) and 3.9 (for baseline items) present the mean RT and log-RT for each L1 group and orthography condition in the matched and mismatched conditions.

Table 3.8. RT (ms) and log-RT data on the test items (minimal pairs and free variants)

Test items		Orth–				Orth+			
		RT		log-RT		RT		log-RT	
L1 Group	Condition	Mean	SD	Mean	SD	Mean	SD	Mean	SD
American	Minimal Pairs	1079	506	3.00	0.17	1183	402	3.05	0.13
	Free Variants	1141	467	3.03	0.15	1083	404	3.01	0.14
Taiwanese	Minimal Pairs	2856	2134	3.35	0.29	1894	1442	3.20	0.24
	Free Variants	2235	1632	3.26	0.26	1920	1322	3.20	0.25

Table 3.9. RT (ms) and log-RT data on the baseline items

Baseline items		Orth–				Orth+			
		RT		log-RT		RT		log-RT	
L1 Group	Condition	Mean	SD	Mean	SD	Mean	SD	Mean	SD
American	Matched	979	484	2.95	0.17	922	297	2.95	0.11
	Mismatched	925	287	2.95	0.12	1066	494	3.00	0.14
Taiwanese	Matched	1304	547	3.09	0.15	1389	748	3.09	0.20
	Mismatched	1446	739	3.12	0.17	1265	603	3.07	0.17

Unlike the accuracy data where the distribution is binomial, the log-RT data is continuous, so a linear mixed-effects model was fitted for the log-RT data. A linear mixed-effects model incorporating the dependent variable log-RT, fixed factors Group (Taiwanese, American), OrthoCond (Orth+, Orth−), Item Type (free variants, minimal pairs and baseline items), Condition (matched vs. mismatched) and FamMinimalPair (to account for an accidental minimal pair found in the familiarization and learning phase) and random factors (subjects and items) was fitted to examine the main effects and possible interactions of the fixed factors on log-RT.

The results showed that there were main effects of L1 group [$F(1, 37) = 42.07$; $p < .001$] and Item Type [$F(2, 14) = 22.37$; $p < .001$] on log-RT, suggesting that Taiwanese participants ($M = 2.97$) were overall significantly slower than Americans ($M = 3.15$) and that the three item types (i.e., minimal pairs, free variants and baseline) caused significantly different RTs: baseline items yielded the fastest RT ($M = 3.001$) and minimal-pair items yielded the slowest ($M = 3.125$). However, there was no main effect of OrthoCond [F(1, 37) = 0.84; $p = .366$] and Condition [F(1, 1474) = 1.34; $p = .247$] on log-RT, but a significant two-way interaction of Group by Item Type [$F(2, 1510) = 21.97$; $p < .001$] and OrthoCond by Item Type [$F(2, 1510) = 3.82$; $p = .022$] where both minimal pairs and free variants yielded faster responses in Orth+ than in Orth–. The three-way interaction of Group by Item Type by OrthoCond [$F(2, 1510) = 8.42$; $p < .001$] was significant, where Taiwanese participants were slower than Americans in both Orth+ and Orth– on the minimal pairs and free variants, but Taiwanese Orth+ were faster on the minimal pairs and free variants than Taiwanese Orth–. A four-way interaction of Group by OrthoCond by Item Type by Condition [$F(3, 1510) = 3.31$, $p = .02$] was also found, which could be due to the lack of mismatched condition for the free variants as well as the matched condition for the minimal pairs.

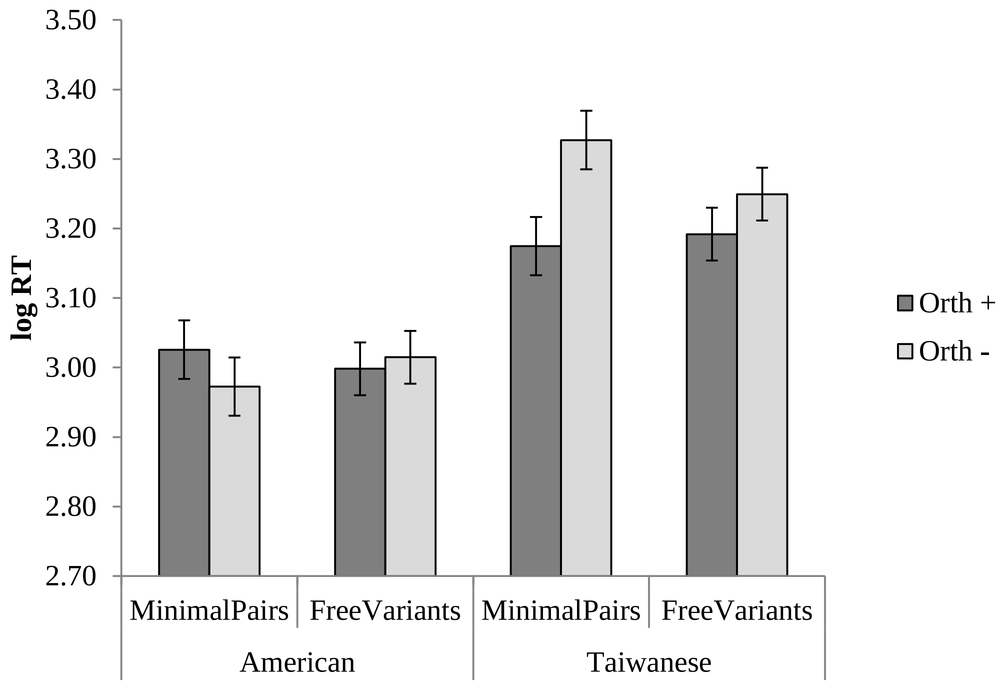


Figure 3.4. Mean log-transformed response times (log-RT) on the free variants and minimal pairs for both L1 groups, as a function of orthographic exposure (dark bars: Orth+; light bars: Orth–)

A post-hoc pairwise comparison using Bonferroni correction for multiple comparisons showed that Taiwanese participants in the Orth+ group responded significantly faster on the minimal pairs than those in the Orth– group ($p = .002$) as shown in Figure 3.4 above, suggesting a possible advantage in processing when orthographic information was available. But such a difference on the minimal pairs was not found for Americans ($p =.269$) (see Figure 3.4 above).

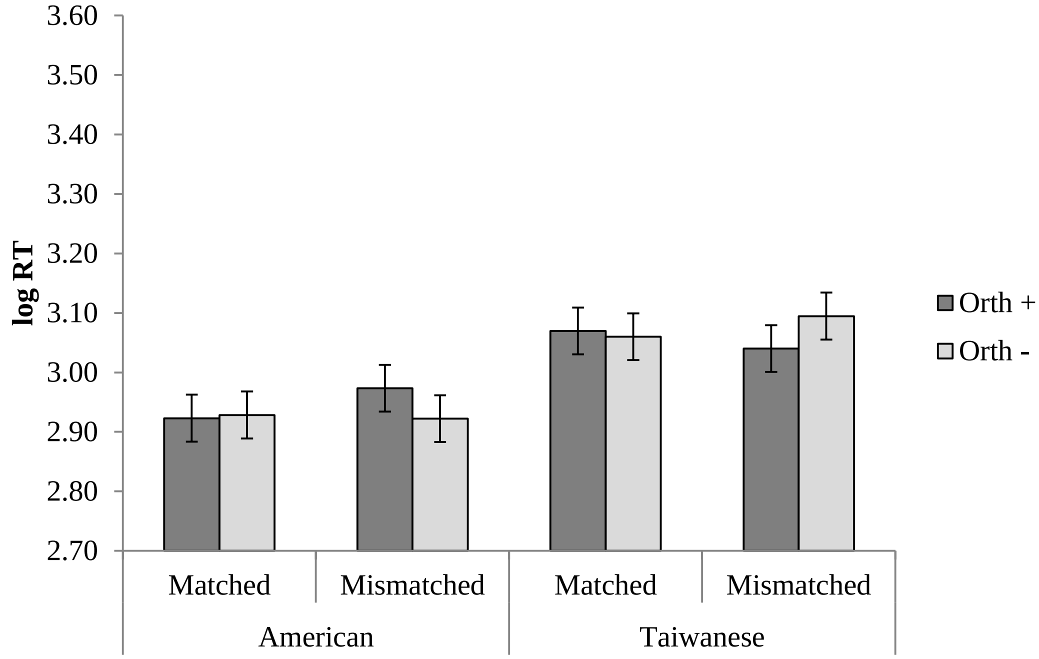


Figure 3.5. Mean log-transformed response times (log-RT) on the baseline items for both L1 groups on the matched and mismatched condition, as a function of orthographic exposure (dark bars: Orth+; light bars: Orth–)

There was no significant difference of log-RT on the baseline items between the two orthographic conditions in both American (p = .600) and Taiwanese (p = .606), suggesting that the processing of the baseline items by both L1 groups was approximately the same (see Figure 3.5 above).

Figure 3.5 above showed that orthographic information did not have a significant facilitative effect on the processing of the baseline items, although Taiwanese Orth– responded more slowly than Taiwanese Orth+.

In addition to the accuracy and RT data, it is crucial to calculate *d'* to examine participants' sensitivity to the vowel alternations as well as the potential response bias. The *d'* data will be reported in the next section.

3.6.4 D-prime data

To calculate the *d'* for each L1 group and orthographic condition, participants' responses were coded as shown in Table 3.10 below.

Table 3.10. Categories of response in picture-word matching

		Vowel alternation		**Baseline**	
		Free Variants	*Minimal Pair*		
Pairing Condition		**Matched**	**Mismatched**	**Matched**	**Mismatched**
Response	**YES**	Hit	False Alarm	Hit	False Alarm
	NO	Miss	Correct Rejection	Miss	Correct Rejection

The *d'* data are shown in Figure 3.6 below.

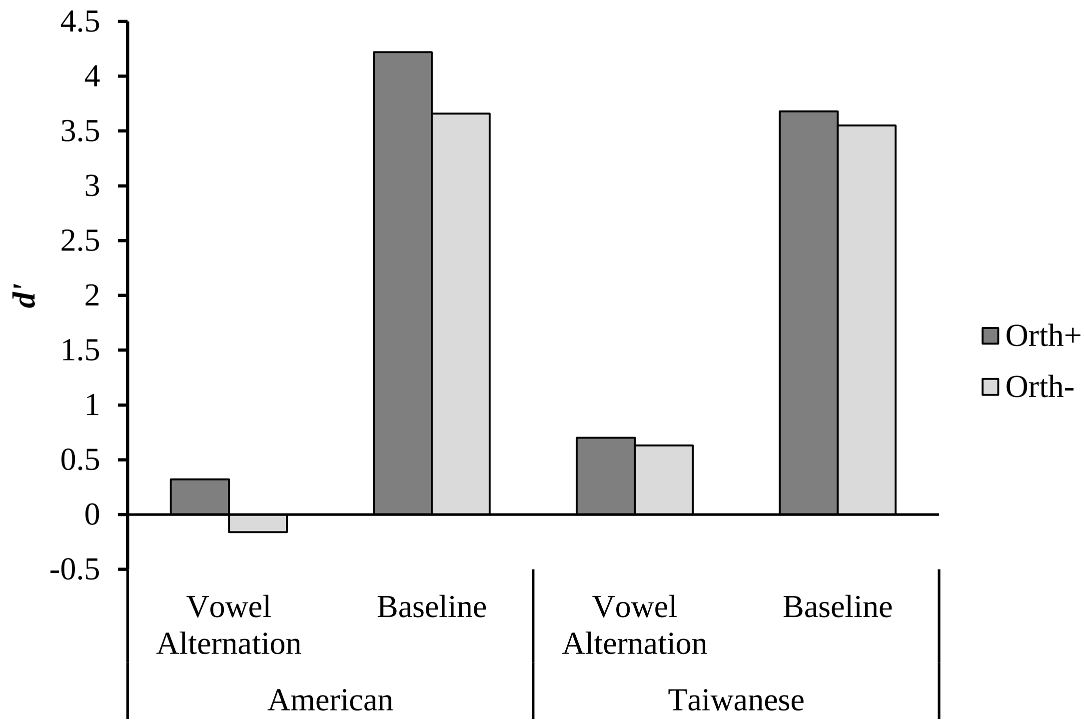


Figure 3.6. Mean *d'* for both L1 groups on vowel alternations and baseline items as a function of orthographic exposure (dark bars: Orth+; light bars: Orth–)

Figure 3.6 above showed that in general the *d'* scores of both L1 groups were slightly higher when orthographic information was available to participants on both the baseline and vowel alternation conditions. The *G* test proposed by Gourevitch and Galanter (1967), which compares the variances of two *d'* scores based on the probability of hits and false alarms, showed that none of the differences between Orth+ and Orth– were significant. Although the differences between Orth+ and Orth– were not statistically significant, orthographic information seemed to benefit Americans more than Taiwanese. The *d'* of American Orth+ was slightly higher than Orth– on the vowel alternations (G = -1.33; G is significant at or above |1.96|) as well as the baseline items (G = -1.07). The *d'* of American Orth– on the vowel alternation was even negative, suggesting that the false alarm rate (i.e., the error rate on the minimal pairs) was higher than the hit rate (i.e., the accuracy on the free variants). For Taiwanese, the *d'* scores on the vowel alternation of both Orth+ and Orth– were about the same (G = -0.2), and they were higher than Americans'. The low *d'* of both L1 groups on the vowel alternations indicates both a high false alarm rate and a high hit rate. In some cases, especially in Americans' detection of vowel alternation, the *d'* even falls below 0, indicating that the false alarm rate is higher than the hit rate. One might wonder whether participants simply responded "*yes*" to most of the items throughout the whole task. However, we can safely reject this possibility, because both L1 groups could detect the matched and mismatched items on the baseline condition very well – they did not just respond "*yes*" to most of the items. Table 3.11 below shows the response bias on each condition. Bias is characterized by criterion *c* where a negative value indicates a "yes" bias while a positive value indicates a bias towards "no".

Table 3.11. Response bias in both L1 groups on both Baseline and Test condition represented in criterion *c*

	Baseline		Test	
	Orth–	Orth+	Orth–	Orth+
American	0.007	-0.285	-1.068	-1.198
Taiwanese	-0.049	0.122	-0.536	-0.627

From Table 3.11 above, Americans showed a strong "*yes*" bias on the Test condition (vowel alternation) in both Orth– and Orth+ group as indicated by the negative *c*. Compared to Americans, Taiwanese showed a smaller "*yes*" bias on the Test condition. Across both L1 groups, participants tended to answer more "*yes*" to the vowel alternating items (i.e., minimal pairs and free variants). However, such response bias was not found on the baseline condition in both L1 groups, although American Orth+ showed a slight bias toward "*yes*". The minimal response bias on the baseline items suggests that participants did not simply respond "*yes*" to most of the items, so the "*yes*" bias found on the Test condition may be attributed to factors other than participants' response strategies.

The *d'* scores showed an opposite trend to the original prediction that orthographic information would benefit Taiwanese more than Americans in the learning of free variation.

Considering that Taiwanese participants' familiarity with the alphabetical writing system may vary as a function of their exposure to English, those who have more exposure to English might utilize the orthographic information differently from those who have less exposure to English. In order to examine the possible effect of their exposure to English on how they use orthographic information in the learning of free variation, two scatter plots (for Orth+ and Orth–, respectively) with length of residence in English-speaking countries (LoR) plotted against *d'* on the test items are shown in Figure 3.7 and 3.8 below. While we might predict such a relationship for the participants in Orth+, it is not expected to play a role for the Orth– group.

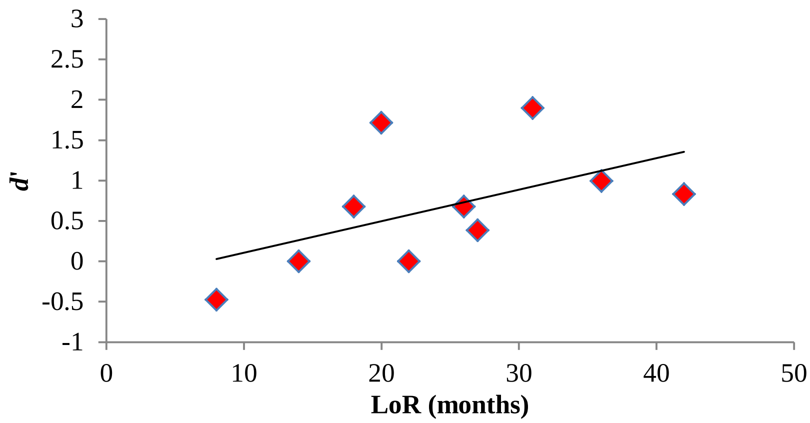


Figure 3.7. Correlation between Taiwanese Orth+ participants' length of residence (LoR) in English-speaking countries and their *d'* on vowel alternations

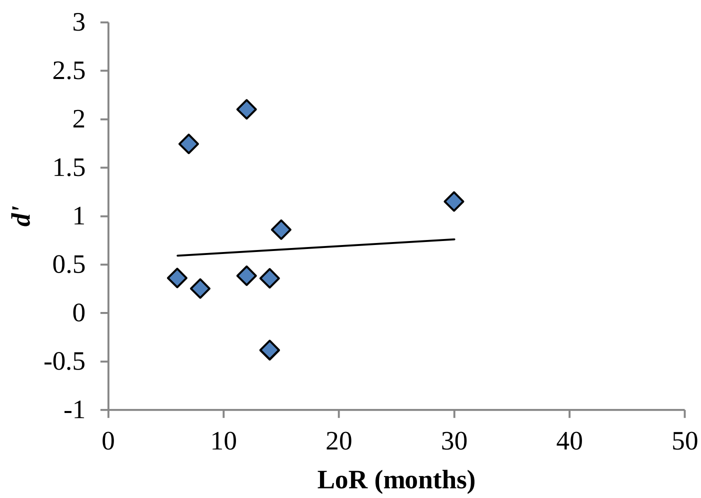


Figure 3.8. Correlation between Taiwanese Orth– participants' length of residence (LoR) in English-speaking countries and their *d'* on vowel alternations

Individual differences can be observed from Figure 3.7 and 3.8 above, and it can be observed that the *d'* score of the Taiwanese Orth+ group tended to be positively correlated with LoR (r = .53; p = .11), while no correlation was found between the *d'* score of Taiwanese Orth– and LoR (r = .09; p = .8). In other words, the longer their length of residence in English-speaking countries, the better their detection of the vowel alternations when orthographic information was available to Taiwanese participants. If the detectability of the vowel alternation was a function of LoR regardless of whether orthographic information was available, there should have been also a positive correlation between *d'* and LoR in Orth–.

3.7 Discussion

To answer each research question and test each hypothesis, the discussion in this section will be organized around each question and hypothesis, and the current results will be compared to previous findings.

Recall that previous studies on the effect of orthography on L2 phonology focus on whether exposure to orthographic forms helps learners distinguish a novel contrast based on the corresponding orthographic forms. The present study, however, investigated whether exposure to orthography helps L2 learners link two free variants to one lexical entry. Although the effect of orthography found in the present study and previous studies taps into two different types of lexical mapping (i.e., one orthographic form to two phonetic variants vs. one orthographic form to one phoneme), a comparison between the findings in the current study and previous studies would provide us a general picture of how orthography influences speech processing.

3.7.1 *RQ1*: Can orthography help L2 learners link two free variants to one single lexical entry?

***H1*: Given the orthographic information during the learning phase, L2 learners would be able to establish a single lexical entry for words with free variants. Learners will be better at associating a picture with two versions of a word differing in one allophone if they are exposed to the spellings than those who do not see the spellings while learning the words.**

The results from Experiment 1 showed that the benefit of orthography on the learning of free variation was limited. Unlike the facilitative effect of orthography on the discrimination of novel phonemic contrasts found in Escudero et al. (2008), in general, there was no significant difference between Orth+ and Orth– on the learning of free variation regardless of L1 background, although Taiwanese responded significantly faster on the minimal pairs when orthographic information was available. Thus, the first hypothesis cannot be fully supported at this point. As in Simon et al. (2010), exposure to orthography during the training phase did not significantly influence L1 English listeners' performance on the categorization of French vowels. As discussed in Chapter 2 (section 2.2.2.3), they attributed this lack of orthographic effect to the fact that the English opaque orthography may make English listeners mistrust the grapheme-to-phoneme correspondence, and thus they might not resort to the orthographic information presented during the training phase. Simon et al.'s explanation may also apply to the findings in this experiment if we consider the orthographic depth of English vowels – the participants in the current study were presumably influenced by their knowledge of English orthography even though the target L2 was not English.

Another possible account for the lack of orthographic effect is the interactive activation between the orthographic and phonological codes. Although the Orth− group did not see the orthographic form of the words during the familiarization and learning phase, we might infer from the Bimodal Interactive Activation Model (BIAM) that they might still mentally form an orthographic representation of the sounds they heard based on English orthography, especially because both L1 groups were experienced alphabetic users and the Americans used alphabet in their L1. However, the grapheme-phoneme correspondence of English vowels is unreliable (e.g., <ea> could be [i] or [ɛ]; <a> could be [e] or [ɑ]), which may cause fuzzy encoding of the target vowels in their memory. For example, when they heard [gekaf] and [gakaf], they might consider these two items as the phonetic realizations of the same word given that both [ɑ] and [e] could both be spelled as <a> in English (e.g., 'dark' and 'bate', respectively). Recall that in the Bimodal Interactive Activation Model (BIAM, see Figure 2.3, section 2.2.2.2), the auditory and visual (orthographic) forms can activate each other. Under this framework, suppose that learners encounter [gakaf] first in the learning phase, for the Orth+ group, seeing <gakaf> might activate [gekaf], [gakaf], [gækəf] or [gakəf], so when they encounter [gekaf] in the testing phase, they might consider that [gekaf] is just another phonetic realization of <gakaf>.

While seeing <gekaf>, the candidates [gekaf], [gekef], [gɛkaf] or [gɛkəf] could enter into the competition to be activated. However, the auditory input of [gekaf] could also activate <gakaf> that s/he encountered earlier, which in turn creates a fuzzy representation of the words. Moreover, the presence of free-variant items may lead Americans to overgeneralize that vowel alternations do not matter in this artificial language.

Several other possible factors that may account for the lack of orthographic effect and the poor learning of free variation are discussed below. They include the grapheme-to-phoneme correspondence of the letter <o>, the differential processing of consonants and vowels, and working memory.

i. The grapheme-phoneme correspondence of the letter <o>.

Recall that in the design, the letter <o> corresponds to [ɔ] and [u]. This grapheme-to-phoneme correspondence is less frequent in English. According to Fry's (2004) revised grapheme-phoneme frequency count originally proposed by Hanna, Hanna, Hodges, and Rudorf (1966), the frequency of <o> - [ɔ] correspondence is 123, whereas the frequency of <o> - [u] correspondence is only 37. Learners will need to familiarize themselves with this less frequent grapheme-phoneme correspondence in the first place, and thus may not be able to quickly link both [u] and [ɔ] to the letter <o>. If the link between [u] and <o> is weaker than that between [ɔ] and <o>, orthographic information in this case would not much help learners in learning that [u] and [ɔ] are free variants, and learners would be less likely to resort to orthographic information to associate the two variants to one single phoneme.

ii. Differential processing of consonants and vowels

Previous studies have shown that consonants and vowels play different roles in lexical access (Bonatti, Peña, Nespor, & Mehler, 2005; Cutler, Sebastián-Gallés, Soler-Vilageliu, & Ooijen, 2000; Luche et al., 2014; New, Araújo, & Nazzi, 2008). For example, Bonatti et al. (2005) found that listeners tend to use the transitional probability (TP) of consonants in word identification more often than the TP of vowels, suggesting that consonants are more closely tied with lexical access.

Cutler et al. (2000) also found that when asked to change nonwords to real words, listeners prefer to change vowels more often than consonants, indicating that vowels tend to constrain lexical selection less tightly than consonants. In other words, consonants play a more important role in lexical access than vowels.

In addition to the above findings from auditory processing, studies on visual processing also reveal similar patterns. New et al. (2008) found that in a visual masked priming lexical decision task, priming effect was found when the consonants of the prime were the same as in the target (e.g., *jalu* → *joli*). However, no priming effect was found when the prime and target shared the same vowels instead of consonants (e.g., *vobi* → *joli*). This finding again suggests that consonants are more closely tied with lexical access than vowels.

In the current study, both L1 groups showed poor detection of the vowel alternations, which is possibly due to the fact that vowels carry lesser weight in lexical access than consonants as found in previous studies. Thus, one possible follow-up study would be to run the same experiment with consonantal alternations, which is reported in Chapter 5.

iii. Memory load

In this task, participants had to learn 42 words in about 40 minutes where 2/3 of the items involved vowel alternations, so the memory load required in this task was relatively high. The task demand could be too high for learners to sufficiently encode the words that differ only in vowels in their memory. As observed in the individual *d'* data, large individual differences were observed, which could also be due to the differences in learners' cognitive abilities.

In a study on the production of unpredictable variation of determiners in an artificial language, Hudson Kam and Chang (2009) found that when task demand was reduced (by providing cues for lexical retrieval), regularization of the variation (i.e., preference for one form over another) was less likely to occur, suggesting that task demand plays a role in regularization. Although their finding was from a study on the learning of morphosyntactic variation, it is also likely that learners might regularize phonological variation from the input, especially when the input is unfamiliar to them and the task demand is high as in the current experiment.

Alternatively, it is also possible that given the high task demand, learners might have not learned the rules of free variation and minimal pairs from the familiarization and learning phase at the first place, and hence there was no generalization happening in the subsequent testing.

3.7.2 ***RQ2*: Does the effect of orthography vary depending on learners' L1 background, specifically, alphabetic vs. non-alphabetic (logographic)?**

***H2*: Taiwanese participants will benefit more from orthography than American learners, because: (a) Chinese and the target language do not share the same orthography, which will result in less interference from L1 orthography; and (b) Chinese learners will rely more on orthographic information in lexical access than users of alphabetic languages as found in studies of visual word recognition.**

Since there was no significant difference between the orthographic conditions in both L1 groups, Taiwanese participants did not benefit from orthography more than Americans. In fact, the *d'* data showed that orthographic information helped Americans detect the vowel alternations slightly better than Taiwanese. Thus, the second hypothesis was also disconfirmed - Taiwanese did not benefit more from orthographic information due to their background of logographic writing system that mismatches with the target language. On the contrary, it is possible that the knowledge of and exposure to the alphabet via their knowledge of English might help.

As shown in the correlation between *d'* and Taiwanese participants' LoR, those who had longer LoR in English-speaking countries detected the vowel alternations better than those with shorter LoR. Taiwanese participants' exposure to alphabetic writing systems can be from various sources (e.g., learning English in schools), their LoR can also be one of the measures for their exposure to alphabetic writing systems. Thus, one interpretation of the positive correlation found in Taiwanese Orth+ is that when they had more experience with an alphabet, they would be better at making use of the orthographic information to facilitate their detection of the vowel alternations. Nevertheless, if we examine the data more closely, it could be found that there was one participant in the Orth– group whose *d'* score was higher than anyone in the Orth+ group, suggesting that s/he could detect the vowel alternations very well even without the orthographic information. Thus, an alternative interpretation of this result cannot be ruled out: instead of facilitating learners' learning, orthographic information might in fact interfere with their learning, such that no one in the Orth+ group had *d'* scores higher than 2. In other words, those who had longer experience with an alphabet might be better at "suppressing" the interference from the orthographic information and detect the vowel alternations better (i.e., longer LoR and higher *d'*) than those who had less experience with alphabet (i.e., shorter LoR and lower *d'*).

Finally, in terms of between-group difference, although the accuracy data showed that Taiwanese participants scored higher than American participants on the minimal pairs, Taiwanese participants also did not score above chance. Motivation is one possible factor that might account for the differences between Taiwanese and American participants. It was observed that during the familiarization and learning phase, the two L1 groups had different learning styles.

For Americans, most of the participants just repeated the word they heard and pressed the button to move on to the next trial. Nevertheless, even with the repeat-and-move on learning style, Americans were still able to learn the baseline items very well. However, for Taiwanese, most of the participants took the task very seriously as if they were preparing for an exam. For example, when they encountered a variant during the familiarization phase, they would stop (pausing during the task was not prohibited), and try to recall the other variant that they had previously seen and learned. By mentally comparing the variants they learned, it was more likely that they were able to detect that the vowel alternations in the minimal pairs changed meanings, which might result in higher accuracy on minimal pairs in the testing phase.

Although the learning of free variation was limited in both L1 groups, the results from the baseline items revealed that learners were still able to learn the words without any vowel alternations. Thus, in terms of the general word learning outcome, both L1 groups exhibited comparable and solid learning performance. Nevertheless, it should be noted that in the current paradigm, the specific learning outcome was measured by a picture-word matching task, and learners learned the words by associating pictures with nonwords of English and Mandarin, which is a type of paired-associate learning. The measure from the picture-word matching task can only test learners' receptive but not productive knowledge. Furthermore, in the current design, the picture-word matching task only involved binary responses instead of multiple choices, so it may even make it easier for learners to decide whether the picture was a correct match with the word than when they had to choose from a set of possible answers. In addition, learners might associate the picture with only parts of the word by remembering some salient cues (e.g., onsets and/or the stressed syllable), making it hard to measure the quality of storage with this kind of binary-response task. Productive knowledge or vocabulary can be tested by a picture naming task which taps the quality of lexical storage and access. Achieving high accuracy on the picture-word matching task does not automatically entail that learners can produce the words accurately or even retrieve the words from their memory. To better understand whether learners have encoded the variants in their memory and can produce it spontaneously, a picture naming experiment was conducted right after the recognition task. The details and results of the picture naming task will be reported in Chapter 4.

Chapter 4. Experiment 2: Production of words with vowel alternations

In Experiment 1 participants could not detect the vowel alternations and learn the free variation very well. Thus, in order to examine the quality and quantity of the encoding and decoding of the lexical items they learned, a picture naming task was conducted in the current experiment. The specific research questions to be answered in this experiment are:

1. Do learners who have exposure to orthographic information produce more correct responses?
2. Do learners encode the free variation in their lexical representation and are they able to utilize the rule of free variation in their production? That is, will they choose to produce the new form of the free variant?

The hypotheses are:

1. Exposure to orthographic information will help learners retrieve and produce the words better than when no orthographic information is available.
2. If learners encode the free variation in memory, it is more likely that they will produce more new forms of the free variant as well as more original forms of the minimal-pair item.

4.1 Materials and Design

The current experiment was conducted right after Experiment 1 to examine learners' quality and quantity of lexical encoding and decoding. Materials for the familiarization and learning phases were the same as those used in Experiment 1. However, in the picture naming task, for the items of minimal pairs, only the pictures from the learning phase were presented, since participants only learned one word from the pair. Each item appeared twice, yielding a total of 32 trials. There were 8 trials of minimal pair items, 8 trials of free variants and 16 baseline trials.

4.2 Participants

Participants were the same as those in Experiment 1: 22 Taiwanese and 23 Americans participated in this experiment. The same participants were dropped from the study as mentioned in Experiment 1. However, due to a technical error in the hardware setting, only 31 participants' responses (out of 40) were collected at the end, of which 20 were females and 11 were males. There were 7 Taiwanese Orth+, 7 Taiwanese Orth–, 9 American Orth+ and 8 American Orth–. The mean age of all participants was 24.78 (*SD* = 5.96), in which the Taiwanese participants' mean age was 30.22 (*SD* = 4.72), and that of the Americans was 20.29 (*SD* = 1.38).

4.3 Procedure

After the testing phase of Experiment 1, participants completed a picture naming task in which they had to name the pictures they learned in the learning phase. In a given trial, a picture was shown in the center of the screen. No spellings were provided.

In this task, a microphone was connected to the E-Prime response box in order to collect the naming latencies. In addition, participants wore a Sennheiser PC350 headset to record their oral responses. As in the previous experiment, E-Prime 2.0 Professional was used for presenting the stimuli as well as collecting the responses. Participants were instructed to name the pictures as quickly as possible and avoid making noises during the task. In addition, if they did not know the answer, they were asked to keep silent and press a key on the response box to skip to the next trial. If they could not remember the whole word, they were also encouraged to produce as many sounds in the word as they could remember.

4.4 Analysis

A total of 992 responses were recorded. Responses that were not clearly recorded were then double-checked in Praat later. No recorded responses were discarded due to noises, false-starts, self-repairs, etc.

Unlike previous studies on the effect of orthography on speech production (e.g., Alario et al., 2007; Damian & Bowers, 2003) where the responses were judged as either correct or incorrect, the analysis for the current experiment employs a more fine-grained scoring scheme by comparing participants' responses with target responses transcribed in the International Phonetic Alphabet (IPA), which also allows us to examine their production of the alternating vowels. Participants' oral responses were scored and coded based on the following criteria, and two types of data were generated:

1. Production score (naming accuracy): The production score (out of 5) was based on segmental errors, for which the scoring criteria are described below:
    1) Consonant errors: for word-final stops, if the place of articulation or voicing was wrong or spirantization occurred (e.g., for a correct response of [p], they produced [t, d, b, g, k, f, v]), half point would be taken off for each error. For stops occurring elsewhere, a voicing error, affrication, or lenition will be given half a point. For fricatives: the errors due to different place of articulation, voicing or fortition (e.g., for a correct response of [f], they produced [s, z, ʃ, ʒ, p,

b]), half point would be taken off for each error. No point would be taken off if the produced segment matched the target segment.

2) Vowel deviations: Except for the free variants and contrasting vowels in the minimal pairs, productions of vowel deviations or vowel reduction would be taken off half point. For example, for the target [e], if participant produced a raised, lowered or reduced vowel (e.g., i, ɪ, ɛ, ə), half point would be taken off.

2. Production of the target free variants and minimal pairs: In order to obtain a more fine-grained measure for the learning of free variation, learners' productions of the target vowel alternations were coded separately as follows.

a.	Produced a new form of the free variant
b.	Produced the original form of the free variant or minimal pairs
c.	Produced a wrong vowel in a word of minimal pairs (e.g., produced [gekaf] instead of [gakaf]

For example, if they learned the variant [fusat] and were able to produce [fɔsat], "a" was coded for this response. However, if participants produced [fusat], "b" was coded for this trial. This coding scheme allows us to differentiate those who could spontaneously produce the new form of the free variant from those who only produced the original form they learned in the learning phase. Note that this coding scheme was *not* used to compute the production score but only for computing the proportion of each type of responses in the production of the "target" alternations.

Also, note that the coding scheme above is categorical rather than continuous, so a generalized estimated equation model on the effect of L1 group (Group hereafter) and Orthographic conditions (OrthoCond hereafter) will be performed on the proportion of production of each category.

It is important to note that the data of naming accuracy (i.e., production score) provides different information than the data of the production of the target alternation. So, for example, for the free variation, if the learned free variant is [fusat] and participant produces [pɔmek], production of new form of free variant ([ɔ]) is coded for this response even though the rest of the sounds are not correct.

The oral responses were first transcribed in IPA by the author. Another phonetically trained transcriber transcribed about 10% of the all responses (four randomly chosen participants' responses) as a measure for transcription reliability. The additional transcriber did not have prior knowledge of the phonetic inventory of this invented language.

The transcriptions done by the additional transcriber were then compared to my own transcriptions. The agreement was measured using the similar scoring scheme described above. The results showed that the two versions of the transcriptions reached 95% agreement, which is deemed a high reliability measure. Note that the purpose of this reliability measure was only to check the reliability of my own transcription, so no reconciliation of the 5% disagreement was performed. For the transcriptions of the 10% responses by the other transcriber, my own version was used for the final scoring.

In addition, as mentioned in section 3.6.2, a free-variant item in the familiarization phase and an item in the learning/testing phase formed a minimal pair by accident. It is possible that this item might yield faster response. The item that formed a minimal pair with another item in the familiarization phase was then coded as 1, and everything else was coded as 0. Thus, a variable FamMinPair (i.e., Familiarization Minimal Pair) was included in the model as a fixed factor to account for this effect.

Self-repairs and false-starts were noted on the scoring sheets. The naming latencies of those trials were later checked in Praat (Boersma & Weenik, 2012). To correct for the voice key's measurement errors caused by self-repairs or false-starts, the naming latencies were determined by the onset of responses in Praat. For the latencies of self-repairs, only the corrected responses were measured.

4.5 Predictions

As mentioned earlier, the purpose of this experiment was to test whether learners could spontaneously produce the words they learned and whether they could spontaneously produce the new forms of the free variants. It was predicted that given the orthographic information during the familiarization and learning phase, the Orth+ group would be able to produce more words and new forms of the free variants, and their production accuracy should be higher than the Orth– group, regardless of the L1 background.

4.6 Results

4.6.1 Production accuracy (score)

Table 4.1 below presents the mean production scores for each group and each orthographic condition. If a participant produced everything correctly, s/he would have a mean score of 5 on each item type.

Table 4.1. Mean production score by Taiwanese and American participants in three items types and in two different orthographic conditions (Orth+ vs. Orth−)

Groups	Item Types	Orth−	Standard deviation	Orth+	Standard deviation
American	MinimalPairs	2.83	1.78	3.76	1.37
	FreeVariants	3.44	1.64	3.64	1.60
	Baseline	3.77	1.58	4.06	1.45
Taiwanese	MinimalPairs	2.79	1.31	2.71	2.03
	FreeVariants	2.84	1.58	3.40	1.56
	Baseline	2.52	1.88	3.47	1.83

A linear mixed-effects model for the production score declaring Group, OrthoCond, Item Type and FamMinPair (i.e., Familiarization Minimal Pair) as fixed factors and Subject and Item as random factors was fitted. In addition, as mentioned in Chapter 3, items that are closer to real English words might yield higher accuracy, so participants' wordlikeness ratings were checked to examine whether they have an effect on production score before fitting the models.

As shown in Experiment 1 (section 3.6.2), the wordlikeness ratings of both the items in the learning phase (henceforth: LearningRating) [$F(2, 14) = 14.43$; $p = .001$] and picture-word matching task (henceforth: TestRating) [$F(2, 13) = 12.23$; $p = .001$] were significantly correlated with Item Types. A model without both LearningRating and TestRating was compared to the model with LearningRating as well as the one with TestRating. The results showed that Item Type was not significant in the models without both ratings [$F(2, 12) = 1.78$; $p = .210$], with LearningRating [$F(2, 14) = 0.475$; $p = .632$], and with TestRating [$F(2, 12) = 0.781$; $p = .479$]. Since Item Type was not a significant factor for production score when wordlikeness ratings were not included in the model in the first place, the correlation between Item Type and wordlikeness rating will not be considered crucial. Thus, the wordlikeness ratings are not included in the final model to simplify model terms.

The linear mixed-effects model showed a main effect of Group [$F(1, 28) = 5.88$; $p = .022$] in which overall Americans ($M = 3.66$) scored higher than Taiwanese ($M = 3.03$), and a significant three-way interaction among Group, Item Type and OrthoCond [$F(2, 940) = 6.19$; $p = .002$] on the production score, showing that American Orth– scored higher on free variants than on minimal pairs, but American Orth+ showed the opposite. Also, Taiwanese Orth+ scored higher on the free variants than minimal pairs but Orth– did not show this pattern. In addition, OrthoCond was marginally significant [$F(1, 28) = 3.44$; $p = .074$] as well as the two-way interaction between Group and Item Type [$F(2, 940) = 2.63$; $p = .072$].

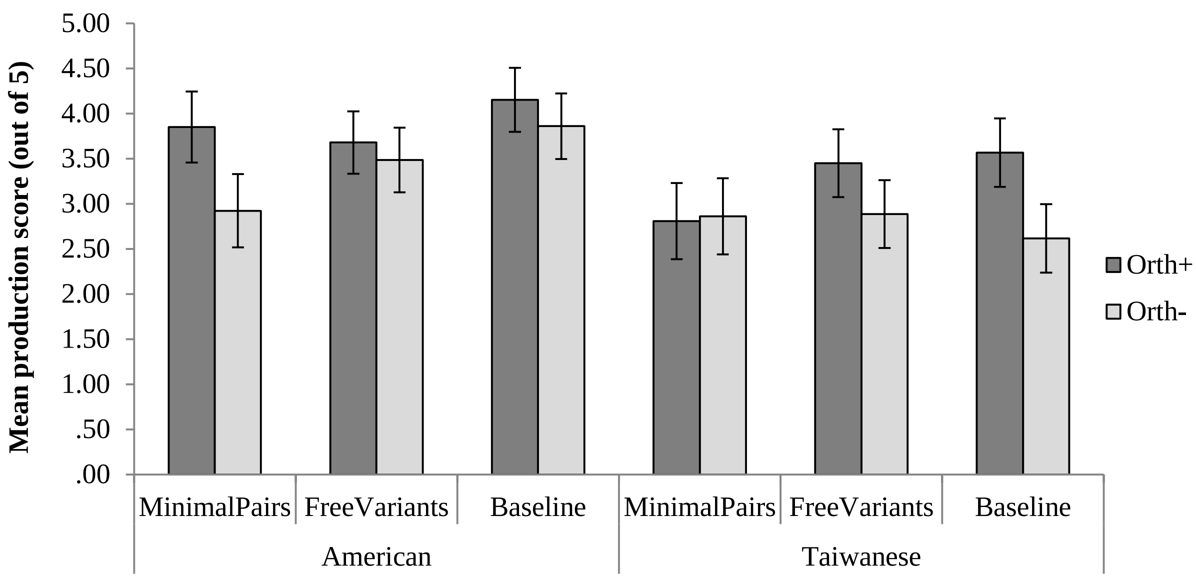


Figure 4.1. Mean production score on three different item types for Taiwanese and American participants[5] as a function of orthographic exposure (dark bars: Orth+; light bars: Orth–)

[5] Graphs are based on the estimated marginal means from the linear mixed-effects model.

Figure 4.1 above demonstrates that Taiwanese participants produced significantly more accurate responses to the baseline items (p = .025) when orthographic information was available to them during the familiarization and learning phase. Although not significant, Taiwanese Orth+ tended to produce more accurate responses to the free variants than Orth–.

American participants, on the other hand, scored significantly higher on the minimal pairs when orthographic information was available (p = .028). For both baseline and free variant items, American Orth+ also tended to score higher than Orth–, although not significantly so.

4.6.2 Naming latencies

Regarding the RT data analysis, the trials where participants did not make a response were dropped from the analysis. An examination of the latency data distribution reveals that the data was positively skewed, so a logarithmic transformation was performed on the data before fitting the statistical models. The log-transformed data is close to normal distribution. Average RTs (ms) for each group and condition are presented in Table 4.2 below.

Table 4.2. Data of naming RT(ms) and log-RT excluding silent trials by Taiwanese and American participants in three items types and in two different learning conditions (Orth+ vs. Orth–)

Groups		Orth–				Orth+			
		RT		log-RT		RT		log-RT	
	Item Types	Mean	SD	Mean	SD	Mean	SD	Mean	SD
American	Minimal Pairs	1808	1327	3.16	0.29	1854	1197	3.19	0.24
	Free Variants	1789	988	3.20	0.21	1797	985	3.20	0.22
	Baseline	1572	887	3.14	0.22	1720	1113	3.16	0.24
Taiwanese	Minimal Pairs	2247	1156	3.29	0.23	2261	1318	3.28	0.26
	Free Variants	2272	1401	3.28	0.25	2122	1257	3.26	0.24
	Baseline	2013	1182	3.24	0.23	2090	1267	3.25	0.23

As in the data of production scores reported in the previous section, a linear mixed-effects model for the logarithmically transformed RT (log- naming latencies) declaring Group, OrthoCond, Item Type and FamMinPair as fixed factors and Subject and Item as random factors was fitted. No main effects of the fixed factors were found, except that L1 Group [$F(1, 28) = 3.91$; $p = .058$] and FamMinPair were marginally significant [$F(1, 12) = 3.5$; $p = .087$]. No interaction was significant.

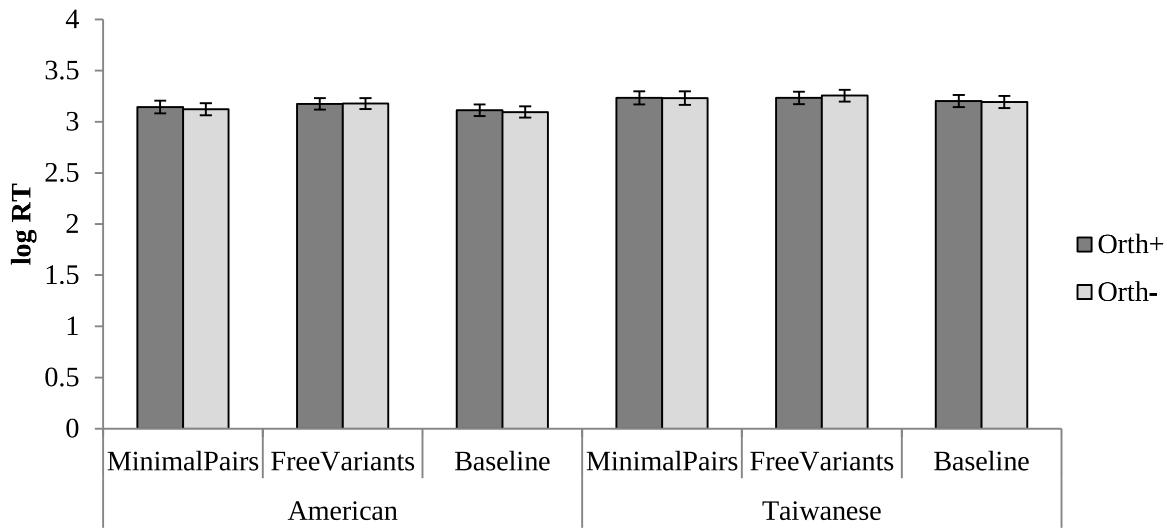


Figure 4.2. Mean log-transformed naming latencies (log-RT) on three different item types for Taiwanese and American participants[6] as a function of orthographic exposure (dark bars: Orth+; light bars: Orth−)

Figure 4.2 above shows that the naming latencies between Orth+ and Orth− were not significantly different across all groups and conditions (all $p > .05$).

4.6.3 Production of vowel alternations

The production scores reported in 4.6.1 cannot demonstrate whether participants produce the new forms of the free variants as well as the correct forms of the minimal pairs. In order to examine participants' productive knowledge of the free variation, the proportion of each response type is reported in Table 4.3 below. The coding was the same as mentioned in section 4.4: a: Produced a new form of the free variant; b: Produced the original form of the free variant or minimal pair; c: Produced a wrong vowel in a word of minimal pairs (e.g., produced [gekaf] instead of [gakaf]).

[6] Graphs were based on the estimated marginal means from the linear mixed-effects model.

Table 4.3. Percentage (%) of the production of vowel alternation by both L1 groups in both orthographic conditions

Group	Orthographic exposure	Wrong minimal pair	Original forms	New form of the free variants
American	Orth+	0.04	0.79	0.18
	Orth–	0.16	0.78	0.06
Taiwanese	Orth+	0.05	0.84	0.11
	Orth–	0.12	0.75	0.12

Note: The data reported here exclude the trials where participants either did not produce the vowel/word or produced a vowel that is beyond the acceptable deviations described earlier. Also, the proportion does not reflect the production accuracy but merely the production of alternating sounds.[7]

Table 4.3 above shows that both L1 groups, when exposed to orthographic forms (Orth+) during the learning phase, produced fewer wrong vowels in the minimal pairs. For those who were not exposed to orthography (Orth–), Taiwanese participants produced equal percentage of wrong minimal pairs and new forms of the free variants (0.12), whereas Americans produced more wrong vowels in the minimal pairs than new forms of the free variants (0.16 vs. 0.06).

Given the categorical data structure, a multinomial generalized estimated equation model was fitted for the production of alternation, which declared Group and OrthoCond as fixed factors and Subject and Item as random factors. Note that Item Type is not a fixed factor in this model, since the category "c" is always associated with the minimal pairs and the category "a" is always associated with the free variants.

[7] The percentages were rounded up to two decimal points, so the combined percentages of the three categories within each subgroup might not be exactly 100% due to rounding up.

The results from the alternation score (i.e., production of the target vowels) showed a main effect of OrthoCond [$\chi^2(1) = 4.65$; $p = .031$], indicating that Orth+ produced fewer wrong minimal pairs than Orth–, but Group [$\chi^2(1) = .186$; $p = .666$] and FamMinPair [$\chi^2(1) = .154$; $p = .695$] were not significant. A significant two-way interaction between Group and OrthoCond was found [$\chi^2(1) = 7.71$; $p = .005$], showing that American Orth– produced more wrong vowels in the minimal pairs than new forms of free variants, but Taiwanese Orth– produced about the same proportion of both types of responses.

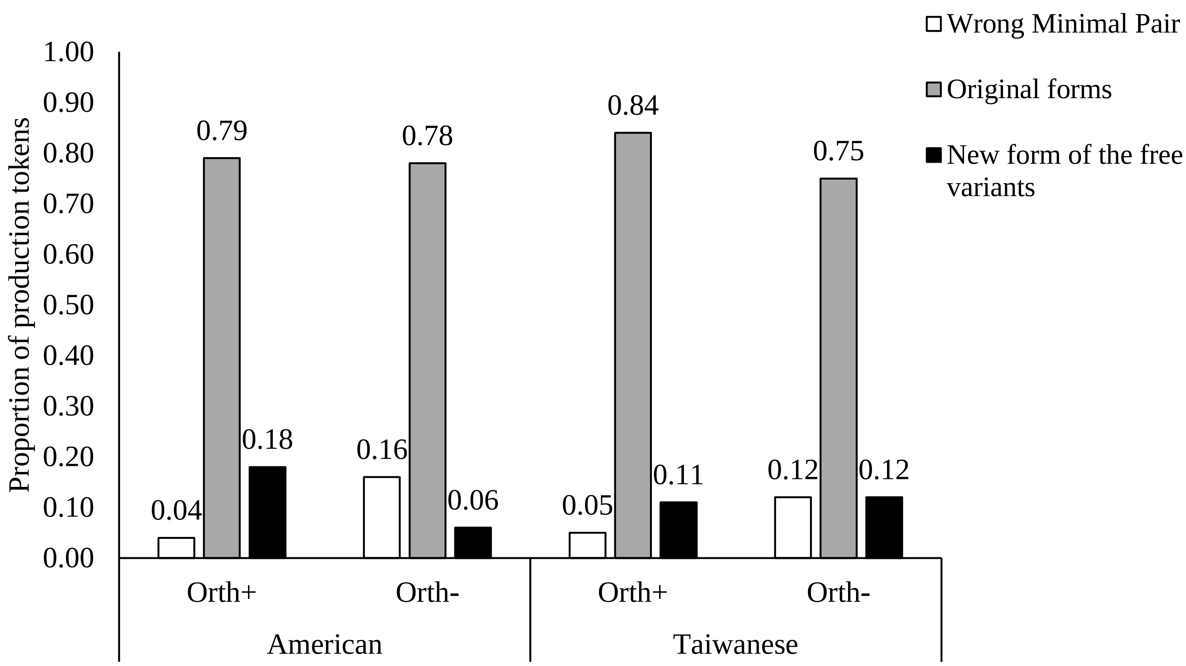


Figure 4.3. Mean percentage for vowel alternation production by Taiwanese and American participants as a function of orthographic exposure (black bars: production of new form of the free variant; grey bars: production of the original forms of both free-variant and minimal-pair items; white bars: production of the wrong vowel in the minimal pair)

Figure 4.3 above demonstrates that both Taiwanese and American learners produced fewer wrong vowels for the minimal pair items when they were exposed to orthographic information during the learning phase. In addition, Americans produced significantly more new forms of the free variants (*M* = 0.18) when orthographic information was available in the learning phase. Comparing the production of the original forms, Taiwanese Orth+ (*M* = 0.84) produced more than Orth− (*M* = 0.75). Overall, Americans benefited more from the orthographic information in producing the new forms of the free variants than Taiwanese.

4.6.4 Comparison between Experiment 1 and 2

In order to compare the results between Experiment 1 and 2, a scatter plot that plots each participant's overall accuracy on the picture-word matching (Experiment 1) against the overall accuracy on picture naming (Experiment 2) across all item types is shown below.

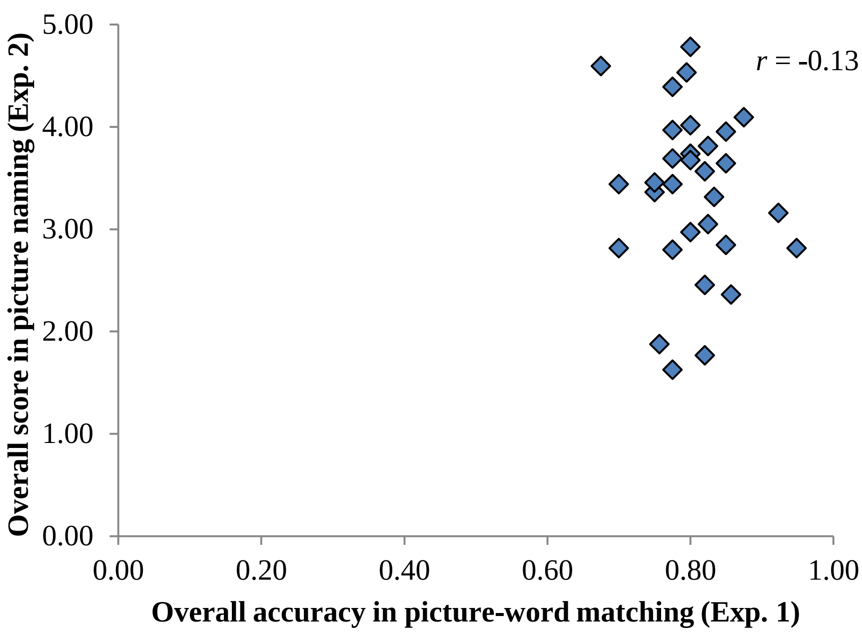


Figure 4.4. Correlation between the accuracy in picture-word matching task (Exp. 1) and picture naming (Exp. 2). Each dot represents each participant's data (subjects 1-12 are not included due to the data loss in Experiment 2).

Figure 4.4 above shows that there is no significant correlation between the overall production accuracy and overall picture-word matching accuracy (r = -0.13), suggesting that high accuracy in the picture-word matching task does not imply high accuracy in the picture naming.

4.7 Discussion

To summarize, the results from the current experiment show that exposure to orthographic forms helped both L1 groups retrieve and produce the words. Most importantly, the significant main effect of orthography in the production of alternations shows that orthographic information helped both L1 groups produce fewer wrong vowels in the minimal pairs as well as helped Americans produce more new forms of the free variants, suggesting that orthographic information is beneficial in decoding the free variation.

Numerous studies have reported mixed results regarding the effect of orthography on word production. Rafat (2011) found that beginning English learners of Spanish showed "non-target-like productions", which could be attributed to their exposure to the orthographic forms during the learning phase. That is, their production of Spanish sounds was influenced by the mismatched grapheme-to-phoneme correspondence between English and Spanish. In Alario et al. (2007), they found that orthography had no effect on the naming of pictures whose names share the first letter even if the grapheme-to-phoneme correspondence is inconsistent. For example, the naming accuracy on the picture whose name begins with <c> is not significantly different from that whose name begins with <k> when both <c> and <k> represent the same phoneme. However, in Damian and Bowers (2003), if participants had to learn words in pairs (i.e., paired-associate learning) before the testing phase where participants had to name the words when cued by the cue words, they were quicker in naming target words which shared the same letter and sound with the cue word, suggesting a benefit of orthography in word naming. In the current experiment, the grapheme-to-phoneme correspondence was also consistent except for the free variants, and both L1 groups generally had higher production scores if they saw the orthographic forms in the learning phase. Thus, the current finding further supports the contention that orthographic information is helpful in word retrieval as found by Damian and Bowers (2003).

A parallel could be drawn between the results from the current experiment and that from a more recent word learning study conducted by Bürki, Spinelli, and Gaskell (2012). In their study, they investigated whether participants would perceive and produce the schwa in the reduced forms of French novel words if they saw the letter <e> after training (e.g., saw <pelour> but heard [pluʁ]). Their results showed that those who saw the spellings with <e> were more inclined to produce the schwa even if the reduced form did not contain the schwa. Notably, this effect of orthography still existed even after a short period of time lapse (i.e., one day after training). They concluded that the exposure to orthography influences how phonetic variant is processed and stored in the mental lexicon, which is consistent with the current findings, so the current results also provide additional support for the benefit of orthography in phonological encoding. The discussion for each research question and hypothesis is presented below.

4.7.1 *RQ1:* Do learners who have exposure to orthographic information produce more correct responses?

***H1:* Exposure to orthographic information will help learners retrieve and produce the words better than when no orthographic information is available.**

The accuracy data from the picture naming task shows that across both L1 groups, Orth+ scored higher than Orth– except for Taiwanese participants on minimal pairs where Orth– and Orth+ were about the same (with a difference of 0.06 out of 5 points). In addition, although Orth+ in general scored higher than Orth–, only American Orth+'s scores on minimal pairs and Taiwanese Orth+'s scores on the baseline items reached statistical significance when examining the production scores by each item type. The reason why only these two item types by specific L1 groups resulted in a significant difference between the orthographic conditions remains unclear at this point, but it is likely that the significant difference found in Taiwanese participants' production score on the baseline items between Orth+ and Orth– reveals a benefit of orthography in word production. Thus, the current results partially support the hypothesis that when orthographic information was made available to learners, they would be more accurate at encoding and spontaneously producing the words. The latency data, however, did not support the hypothesis, as there was no effect of orthography in either L1 group. As mentioned earlier, the effect of orthography was found on the production of phonetic variants in Bürki et al. (2012). One major difference between the present study and their study is that their participants were asked to write down the spellings when they saw them during learning, but in the current design participants were not informed of the presentation of spellings during the familiarization and learning phase. In other words, participants in the Orth+ group were not required to use the orthographic information in their learning. Indeed, the effect of orthography could be stronger if learners were required to write out the words.

As mentioned above, the effect of orthography is most significant on the baseline items produced by Taiwanese participants and the minimal pairs produced by Americans. Considering that the production score only reflects the general production accuracy of each item type and does not reflect how well participants produced the vowel alternations, it is necessary to examine the data of the alternation scores to discuss the effects found in Americans' production of the minimal pairs.

4.7.2 ***RQ2*: Do learners encode the free variation in their lexical representation and are they able to utilize the rule of free variation in their production? That is, will they choose to produce the new form of the free variant?**

***H2*: If learners encode the free variation in their memory, it is more likely that they will produce the new form of the free variant and produce the original form of the minimal-pair item.**

Hypothesis 2 was only partially confirmed. In general, as shown in Figure 4.3 above, American learners show a clear benefit of the orthographic information in the production of vowel alternations: American Orth+ produced a significantly lower percentage of wrong vowels in the minimal pairs but higher percentage of the new forms of free variants than Orth–. Note that the production of the free variants alone cannot be completely attributed to participants' learning of the free variation. If one produces a free variant, it is also possible that they are primed by the items shown in the picture-word matching task. However, the production of the contrasting vowels in the minimal pairs provided another clue of participants' learning of free variation, which complements the results of the free variant production.

Nevertheless, the effect was not strong in Taiwanese participants, although the difference between Orth+ and Orth– is still observable: the Taiwanese Orth+ tended to produce fewer wrong vowels of the minimal pairs than Orth–, suggesting that exposure to orthographic forms helped both L1 groups learn that /e/ and /a/ are contrastive. But such a difference was not clear in Taiwanese participants' production of the new forms of the free variants where Orth+ and Orth– produced about the same proportion.

Another caveat is that the proportion of the original forms of both the minimal pairs and free variants is at least 75% in both L1 groups and both orthographic conditions, indicating participants' preference for the forms they learned in the learning phase, which in turn suggests a small degree of the free variation learning. Given this caveat, the results from picture naming still provide additional information regarding the learning of free variation. Recall that in Experiment 1, participants showed poor detection of the vowel alternations (low *d'*), especially American participants who showed a bias towards "*yes*" and had low accuracy on the minimal pairs. In the picture naming task, however, participants were able to produce fewer wrong vowels in the minimal pairs, especially Orth+, showing a pattern opposite to the results in Experiment 1. As shown in the null correlation between the overall accuracy in Experiment 1 and 2, the performance on the picture-word matching does not relate to the performance on the picture naming, and vice versa.

Based on the results from Experiment 1 and 2, a firm conclusion on whether orthographic forms can help learners learn a free variation cannot be made because of the poor detection of the vowel alternations in Experiment 1 and the small proportion of the new form of free variant production in Experiment 2. As discussed in Chapter 3, the differential roles of vowels and consonants in lexical access might contribute to this poor detection of the vowel alternation. Thus, in the next experiment, I will investigate whether learners can learn to associate free variants with one lexical entry more accurately when this involves consonantal alternations.

Chapter 5. Experiment 3: Learning of consonantal free variation

In Chapter 3 (Experiment 1), several possible factors with regard to the participants' poor performance on the detection of the vowel alternations were discussed. Two of the possible accounts are the different roles of consonants and vowels in lexical access and the inconsistent grapheme-to-phoneme correspondence of English vowels. First, as discussed in Chapter 3 (section 3.7.1), several studies have found a potential difference between consonants and vowels regarding their role in lexical access (e.g., Bonatti, et al., 2005; Cutler et al., 2000) where vowel variation may lead to less accurate word recognition than consonants. Second, the inconsistent grapheme-to-phoneme correspondence in English might make the orthographic information less useful for learners to associate the free variants to one lexical entry. Given the potential difference between consonants and vowels in terms of their roles in lexical access as well as grapheme-to-phoneme correspondence in English, the current experiment replicates Experiment 1 with new items that involve consonantal alternations to examine whether participants can learn to associate the consonantal free variants with one lexical entry more accurately. The specific research questions to be answered in this experiment are:

1. Can learners learn to associate the consonantal free variants with one lexical entry more accurately than the vocalic one as shown in Experiment 1?
2. Will orthography benefit the learning of free variation when it involves consonants?

   2.1 If the answer to Question 2 above is "yes", will there be a between-group (i.e., Taiwanese vs. American) difference in the effect of orthography and the learning of free variation? Although an effect of L1 group was not found in Experiment 1, it is still questionable whether learners' L1 background (alphabetic vs.

logographic) will influence learners' reliance on the orthographic information in word learning.

The hypotheses to be tested are:

1. If consonants are tied with lexical access more than vowels as found in previous studies, it is hypothesized that learners would be able to learn the free variation better than vowel alternations.
2. Because the grapheme-to-phoneme correspondence of consonants in English is more transparent than that of vowels, learners will benefit more from the orthographic information in detecting the consonantal alternations and learning the free variation.
   2.1 Since the target language's grapheme-to-phoneme correspondence of the consonants matches with that of English, American participants would show a larger effect of orthography than Taiwanese participants.

5.1 Design and Paradigm

The design and the paradigm of the current experiment is the same as that of Experiment 1 (see section 3.1)

5.2 Participants

Twenty-three Taiwanese and twenty-seven Americans were recruited for this experiment. The participants were different from those in Experiments 1 and 2. They were paid $10 for their participation regardless of the amount of time they spent or their performance. None of the participants were linguistics-related majors/minors, although a few of them took a course or two in introductory linguistics in non-linguistics programs. Among the 50 participants, nine of them were dropped from the final analysis for reasons described below, involving their language and demographic background. One American participant reported three concussions occurred at different times caused by injury when playing sports within the last 5 years. She also noted a short-period of memory loss following the concussions but stated that it did not interfere with her learning after recovery. The result of the baseline items showed that she scored 96% correct, which was comparable to other participants. However, considering her history of short-period memory loss and unknown brain regions of lesion, which could also potentially affect phonological processing, and that the injury occurred within less than 5 years, her data was excluded from the analysis. Another American's data was excluded because his accuracy on the AX discrimination fell below 90%. In addition, one American who needed more than 4 learning cycles did not complete the experiment due to the excessive amount of time needed to complete. Another five participants' data (all in the Orth− group, and these participants also did the same version of the stimuli list) were discarded due to scripting error. One Taiwanese participant reported that she was diagnosed of some degree of hearing problems, and she could hardly hear any tones in the hearing screening or name any picture (to be reported in Experiment 4 next) in the first attempt. As a result, her data was also excluded from the analysis. In the end, the data of 41 participants (29 females and 12 males) were kept for final analysis, of which 20 were Taiwanese (mean age = 26.5, *SD* = 3.14; mean LoR in English-speaking countries = 12 months, *SD* = 8.97) and 21 were Americans (mean age = 20.92, *SD* = 1.17).

Three other participants reported minor problems: One Taiwanese reported one instance of mild concussion during childhood and had fully recovered since then, so her data was included. Two American participants reported mild speech disorder in childhood (difficulties in producing [ɹ]) but had been corrected after speech therapy, so their data were still included.

5.3 Materials

As in Experiment 1, 54 novel, word-like nonwords were created (see Appendix II), and audio-recorded for presentation in the picture matching experiment. The same 40 black-and-white line drawings were selected from the same picture database used in Experiment 1 (Alario & Ferrand, 1999) to pair with the nonwords.

5.3.1 Orthographic depth

As in Experiments 1 and 2, in order to make the invented language a transparent one and avoid possible interference from participants' opaque L1, the grapheme-to-phoneme correspondence in the invented language was kept consistent and similar to that in Experiment 1. A speaker from a transparent language was selected to record the stimuli. Thus, a German near-native speaker was chosen to record the materials. The grapheme-to-phoneme correspondence (see Table 5.1 below) was similar to that in Experiment 1 except for the correspondence of <t> and <p>, which are used for free variants, and that of <o>, which corresponds to [ɔ] only rather than to both [ɔ] and [u] used in Experiment 1.

Table 5.1. Grapheme-to-phoneme correspondence in the consonantal experiment (Exp. 3)

Sounds	**Letters**
[t], [d]	<t> (when [t][d] are free variants) or <t> and <d>, respectively (when [t][d] are contrastive)
[p], [b]	<p> (when [p][b] are free variants) or <p> and <b>, respectively (when [p][b] are contrastive)
Consonants other than above	One-to-one correspondence (e.g., [k] only corresponds to <k> but not <c>)
[ɪ], [ɛ]	<i>, <e> (in unstressed syllables)
[i], [e]	<i>, <e> (in stressed syllables)
[a]	<a>
[ɔ]	<o>
[u]	<u>

5.3.2 Stimuli recording

As in Experiment 1, one single talker was recruited to record the stimuli. The same German near-native female speaker in Experiment 1 recorded all the stimuli for the current experiment. Stimuli were recorded in a quiet room using Marantz PMD 620 digital portable recorder and spliced using Adobe Audition 5.5. The amplitude of all stimuli was matched to −30.5 dB total RMS.

5.3.3 Constructing stimuli

Similar to Experiment 1, all stimuli were disyllabic nonwords (CV.CVC) based on the ARC nonword database (Rastle et al., 2002) and MRC Psycholinguistic database (Coltheart, 1981). All stimuli were first-syllable stressed and modified from Experiment 1, but the constraints of the stimuli were kept the same as in Experiment 1.

As in Experiment 1, there were three types of items: free variants, minimal pairs, and the baseline items (i.e., words that did not involve any alternations). The number of items of each condition in each phase was the same as in Experiment 1.

The alternations always occurred on the word-initial position as well as the stressed syllable, so that the results from this experiment could be comparable to that of Experiment 1 where the vowel alternations also occurred on the stressed syllable, which is more perceptually salient than the unstressed syllables (e.g., Carroll & Shea, 2007).

None of the stimuli were real words in participants' native languages (i.e., Mandarin and English). As mentioned in section 3.3.3, to control for the possible effect of neighborhood density, the English phonological and orthographic neighborhood density of the nonwords in the current experiment were obtained from the Speech and Hearing Lab Neighborhood Database of

Washington University in St. Louis (Sommers, 2003), which is based on the Hoosier Mental Lexicon (Nusbaum et al., 1984). The average number of orthographic neighbors is 0.15 (*SD* = 0.36) and that of the phonological neighbors is 0.04 (*SD* = 0.19). In addition, the proportion of different categories of consonants and vowels was approximately evenly distributed in each position.

5.3.4 Choosing the consonantal alternations

The consonantal alternations are selected based on the English consonant inventory, since both groups of participants were either English native speakers or advanced learners of English (i.e., American and Taiwanese) and were familiar with the English phonemic inventory. As mentioned in Experiment 1, a pair that is perceptually distinguishable is needed, because: 1) if learners cannot distinguish the contrast in the first place, it will be hard to know whether they have learned that the two variants map to the same lexical entry (i.e., free variation). For example, if they cannot tell whether A sound is different from B sound, their mapping of both A and B to the same lexical entry cannot be interpreted to mean that they have learned the free variation but that they do not know A and B are different; and 2) unlike the study by Showalter and Hayes-Harb (2013) where learners were learning a novel contrast, the present experiment examines whether learners can associate two "distinguishable" variants to one lexical entry. The pairs [p^h][b] (broad transcription [p][b] hereafter) and [t^h][d] (broad transcription [t][d] hereafter) were chosen for this study. Although Mandarin does not have a voiced stop, the contrast between short and long lag stop in Mandarin (e.g., [p] vs. [p^h]; [t] vs [t^h]) (e.g., Chao and Chen, 2008) is functionally equivalent to the stop voicing contrast in English. In other words, both Mandarin and English employ a two-category VOT distinction (Keating, 1981). Furthermore, in American

English the voiced stops /b//d//g/ in syllable-initial position are usually realized as unaspirated voiceless stops [p][t][k] (Ladefoged, 2000: 57). Thus, both Mandarin and English native listeners should be able to clearly distinguish both pairs, as is supported by the results from the AX discrimination task where the mean accuracy was 0.98 (*SD* = 0.13), suggesting that participants could clearly distinguish the target pairs of consonants.

Two lists of items were created. In the first list the [p-b] alternation was free variation, while the /t-d/ alternation was contrastive. In the second list the conditions were reversed: the /p-b/ alternation was contrastive, while the [t-d] alternation was free variation. The conditions between subjects were counterbalanced: half of the participants received the first list, and the other half received the second list. In addition, half of the free-variation and minimal-pair items in the learning phase were Variant A and Word A, respectively, while another half of the items in the learning phase were Variant B and Word B, respectively. For example, in the case of [p-b] free variation, two of the free variants begin with [p], and another two begin with [b] in the learning phase. Then for those beginning with [p] in the learning phase, they will appear as [b] in the testing phase, and vice versa.

5.4 Procedure

The word learning procedures were the same as in Experiment 1, with the exception that one half of the participants learned [t-d] as free variants while the other half learned [p-b] as free variants.

At the end of the experiment, as in Experiment 1, they were asked to rate the wordlikeness of each item used in the experiment. However, considering that the instruction of the wordlikeness rating used in Experiment 1 was somewhat confusing for some participants, in the current experiment, participants were asked whether the pronunciation of the word reminded them of the pronunciation of another real word in any language they know. They were instructed to rate on a Likert scale from 1 to 5 (5 if the pronunciation is very close to a real word and 1 if it is totally unlike a real word).

5.5 Predictions

1. The accuracy on the baseline items should be similar to that in Experiment 1 regardless of L1 background.
2. A significant effect of orthography will be observed in both L1 groups, considering that the letter-sound correspondence of consonants is more reliable than that of vowels.
3. Taiwanese participants might show slight improvement in the accuracy on minimal pairs, whereas Americans would show significant increase in the accuracy on minimal pairs. In other words, Americans will learn the free variation more accurately (as shown by the increased accuracy on minimal pairs) when the alternations are consonantal.

5.6 Results

The data presentation in this experiment is similar to that of Chapter 3 (Experiment 1). In section 5.6.2 as well as 5.6.3 below, I will first present the observed means and standard deviations in tables followed by bar graphs where the means are estimated marginal means from the statistical models.

5.6.1 Data structure and preparation

As in Experiment 1, the first block of the testing phase contains the crucial experimental condition in the current study, so that both accuracy and response time (RT thereafter) data in the first block are analyzed. RT data were logarithmically transformed before running analysis of variance due to its skewed distribution.

Similar to Experiment 1, the independent variables include: L1 groups (Group hereafter), Item Type (i.e., free variants, minimal pairs, and baseline items), pairing conditions (i.e., matched vs. mismatched) (Condition hereafter), orthographic conditions (i.e., Orth+ vs. Orth–) (OrthoCond hereafter) with the addition of Contrast Type (i.e., /t-d/ vs /p-b/). As mentioned in Chapter 3, each participant's learning ability may vary, so one item type might be easier for certain subjects than the other. In other words, some people might be slower on certain items due to their learning ability or other cognitive factors. Therefore, as in Experiment 1, to account for this, the data was trimmed by standardized residuals from Subjects and Item Type: A univariate analysis of variance (ANOVA) with logarithmically transformed RT (log-RT hereafter) as the dependent variable and Subjects and Item Type (free variant, minimal pair and baseline) as fixed factors was run to obtain the standardized residuals (*Z* residuals) of the log-RT data. In order to include more data, instead of trimming by 2 standardized residuals (which would leave about 95% of all data points), the trials above or below 3 standardized residuals were trimmed. In the end 18 out of 1640 trials (1.1%) were trimmed.

5.6.2 Accuracy data

Table 5.2 below shows the mean accuracy on minimal pairs and free variants across both contrast types by both American and Taiwanese participants.

Table 5.2. Overall mean accuracy (across both contrast types) on the consonantal test items (minimal pairs and free variants)

Test items		Orth–		Orth+	
L1 Group	Condition	Mean	SD	Mean	SD
American	Minimal Pairs	0.66	0.47	0.86	0.34
	Free Variants	0.46	0.50	0.55	0.50
Taiwanese	Minimal Pairs	0.49	0.50	0.54	0.50
	Free Variants	0.39	0.49	0.55	0.50

Table 5.3 below shows the mean accuracy on minimal pairs and free variants by both American and Taiwanese participants for each contrast type.

Table 5.3. Mean accuracy on the test items for each contrast type

Contrast Type		[p-b] free variation				[t-d] free variation			
Test items		Orth–		Orth+		Orth–		Orth+	
L1 Group	Condition	Mean	SD	Mean	SD	Mean	SD	Mean	SD
American	Minimal Pairs	0.60	0.49	0.94	0.24	0.73	0.45	0.78	0.42
	Free Variants	0.43	0.49	0.62	0.49	0.50	0.50	0.48	0.50
Taiwanese	Minimal Pairs	0.61	0.49	0.55	0.50	0.38	0.49	0.54	0.50
	Free Variants	0.25	0.43	0.55	0.50	0.53	0.50	0.55	0.50

Table 5.4 below shows the mean accuracy on the baseline items (regardless of contrast types) by both American and Taiwanese participants in both matched and mismatched pairings and in both orthographic conditions.

Table 5.4. Mean accuracy on the baseline items

Baseline items		Orth–		Orth+	
L1 Group	Condition	Mean	SD	Mean	SD
American	Matched	0.93	0.25	0.95	0.22
	Mismatched	0.95	0.22	0.95	0.22
Taiwanese	Matched	0.93	0.25	0.92	0.26
	Mismatched	0.87	0.34	0.93	0.25

Before analyzing the accuracy data, it is necessary to check whether the two groups had significantly different wordlikeness ratings and whether a particular item type is more like a word than another, because, as mentioned in Chapter 3, a more word-like item might yield higher accuracy. A linear mixed-effects model for wordlikeness ratings with Group and Item Type as fixed factors and Subject and Item as random factors showed that, unlike in Experiment 1 where wordlikeness ratings were significantly correlated with Item Type, wordlikeness ratings in Experiment 3 were not correlated with either Item Types or Group (both $p > .1$). In other words, the ratings were not significantly different among Item Types as well as between the two L1 groups, so the wordlikeness ratings were not included in the models of accuracy, RT and naming accuracy.

Note that there were two items (baseline item “nekis” and test item “dumak”) that received higher wordlikeness ratings (3.8 and 4 out of 5, respectively) by Taiwanese participants. All other items were rated below 3.5. However, these two items were not responded to more accurately than other items from the same Item Type. “nekis” yielded 95% correct, but another item (“liguk”) rated 2.2 yielded 98% accuracy. And “dumak” yielded 55% correct, while another item “tamik” rated 2.8 yielded 65% accuracy. In addition, when wordlikeness ratings were included in the generalized estimated equation model for accuracy as a fixed factor, the wordlikeness ratings of the learning-phase items and that of the testing-phase items were both not significant [Learning-phase items: $\chi^2(1) = 1.46$, $p = .226$; Testing-phase items: $\chi^2(3) = .358$; $p = .550$]. Therefore, in the final model the wordlikness ratings were excluded.

A generalized estimated equation model for the accuracy data declaring Subjects and Item as random factors and the orthographic conditions (OrthoCond), L1 groups (Group), pairing condition (Condition), the types of items (Item Type) and Contrast Type (/b-p/ or /t-d/) as the fixed factors was fitted. The factor Contrast Type is to examine whether different place of articulation would have an effect on the learning of free variation.

The generalized estimated equation model for the accuracy data showed a main effect of Group [$\chi^2(1) = 13.16$; $p < .001$] where Americans (*M*[8] = 0.81) were overall more accurate than Taiwanese (*M* = 0.69), and a main effect of Item Type [$\chi^2(2) = 144.2$; $p < .001$] where the baseline items yielded the highest accuracy rate (*M* = 0.94), minimal pairs the second highest (*M* = 0.70) and the free variants the lowest (*M* = 0.45). Lastly and most importantly, the main effect of OrthoCond [$\chi^2(1) = 7.83$; $p = .005$] indicated that generally Orth+ (*M* = 0.80) outperformed Orth– (*M* = 0.71). Contrast Type [$\chi^2(1) = .420$; $p = .517$] and Condition [$\chi^2(1) = .594$; $p = .441$] were not significant. The results also showed a significant three-way interaction of Group by Item Type by Contrast Type [$\chi^2(2) = 6.6$; $p = .037$] where across both orthographic conditions, Americans scored higher on the free-variant items than Taiwanese when [p-b] were free variants but not when [t-d] were free variants, and Taiwanese scored higher on minimal-pair than free-variant items when [p-b] were in free variation but not when [t-d] were in free variation. Another three-way interaction of Group by Contrast Type by Condition [$\chi^2(1) = 4.74$; $p = .030$] was also

[8] The means reported in the inferential statistics are based on the estimated marginal means from the linear mixed effects model, while the means reported in the tables are the observed means.

found, in which Americans scored higher on matched condition than mismatched when [t-d] were free variants but not when [p-b] were free variants. A marginally significant interaction of Group by OrthoCond by Contrast Type [$\chi^2(1) = 3.31$; $p = .069$] was also found. No other interactions were found.

I now turn to the analysis by contrast type, to see whether the two contrasts have a different effect on the learning of free variation. Figure 5.1 and 5.2 present the mean proportion correct for each contrast and each group; Figure 5.3 presents the mean proportion correct for the baseline items. As can been seen in Figure 5.1 below, in general participants in the Orth+ condition scored significantly higher than those in the Orth– condition as manifested by the main effect of Orthographic Condition, except for Taiwanese participants' accuracy on the minimal-pair items.

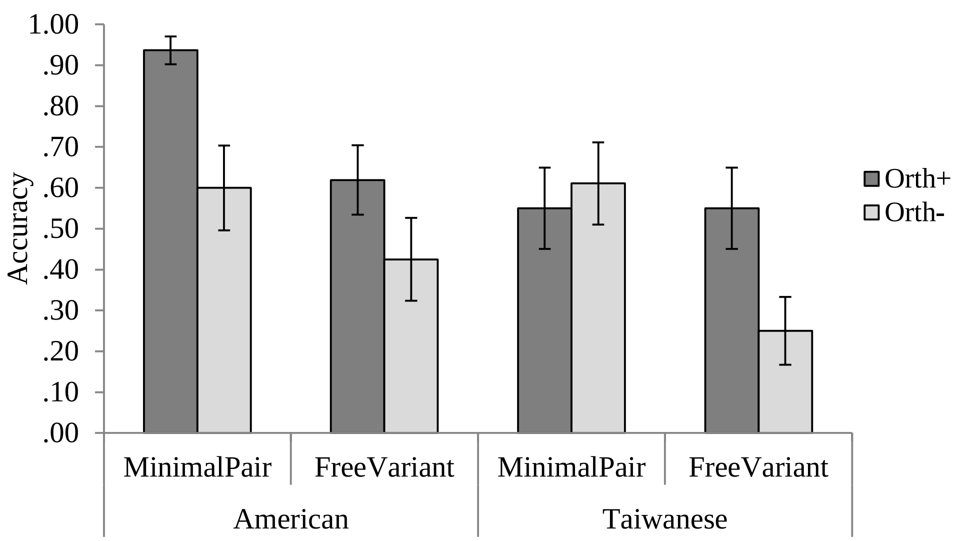


Figure 5.1. Mean proportion correct on the free variants [p-b] and minimal pairs /t-d/ for both L1 groups, as a function of orthographic exposure (dark bars: Orth+; light bars: Orth–)

A post-hoc pairwise comparison with Bonferroni correction for multiple comparisons showed that when [p-b] were in free variation, American Orth+ scored significantly higher than Orth– (p = .002) on the minimal pairs, and Taiwanese Orth+ scored significantly higher than Orth– on the free variants (p = .02) (see Figure 5.1 above).

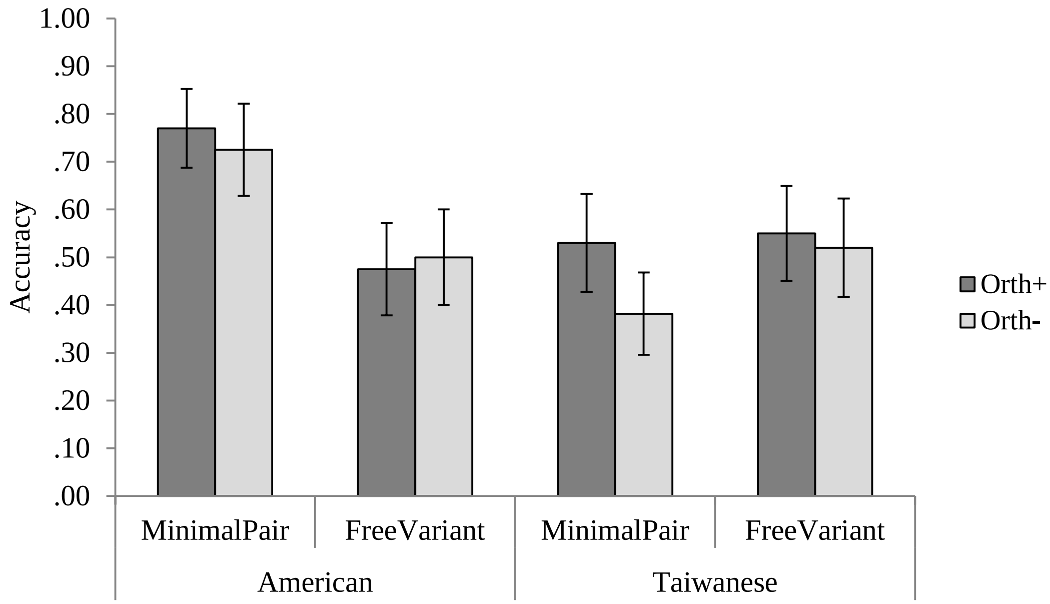


Figure 5.2. Mean proportion correct on the free variants [t-d] and minimal pairs /p-b/ for both L1 groups, as a function of orthographic exposure (dark bars: Orth+; light bars: Orth–)

Figure 5.2 above showed that, unlike the previous case where [p-b] was in free variation, participants in the Orth+ condition did not score significant higher than those in the Orth– condition (all p values > .1).

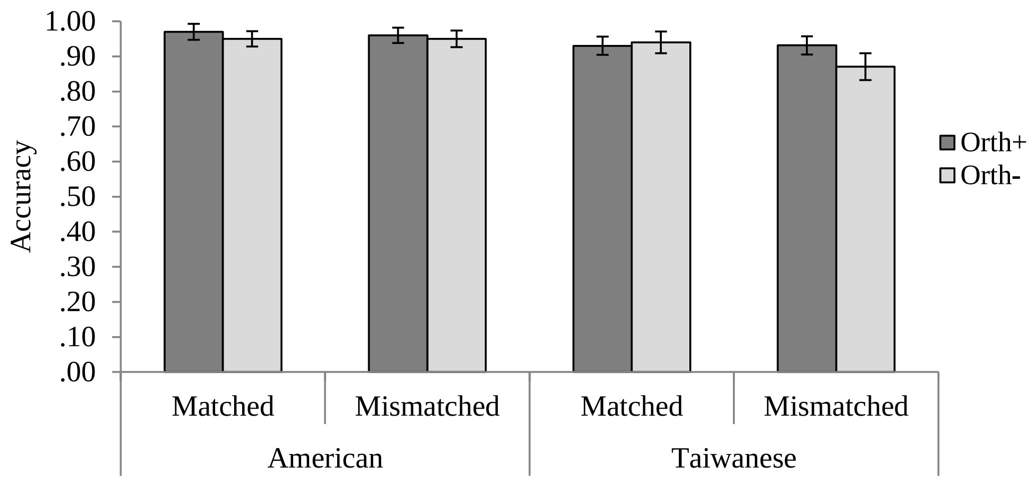


Figure 5.3. Mean proportion correct on the baseline items for both L1 groups, as a function of orthographic exposure (dark bars: Orth+; light bars: Orth–).

In terms of the baseline items, Figure 5.3 above shows that there was no significant difference in accuracy between Orth+ and Orth– in both L1 groups, although Mandarin Orth+ scored slightly higher than Orth– on the mismatched condition (p = .117).

5.6.3 RT data

Table 5.5 below shows the mean RT and log-RT on minimal pairs and free variants across both contrast types by both American and Taiwanese participants.

Table 5.5. Overall mean RT (ms) and log-RT on the test items (minimal pairs and free variants, across both contrast types)

Test items		Orth–				Orth+			
		RT		log-RT		RT		log-RT	
L1 Group	Condition	Mean	SD	Mean	SD	Mean	SD	Mean	SD
American	Minimal Pairs	1194	453	3.05	0.14	1187	404	3.05	0.13
	Free Variants	1061	315	3.01	0.13	1237	510	3.07	0.14
Taiwanese	Minimal Pairs	1725	1145	3.17	0.23	1691	800	3.19	0.19
	Free Variants	1815	1003	3.19	0.25	1799	1163	3.19	0.24

Table 5.6 below shows the mean RT on minimal pairs and free variants by both American and Taiwanese participants when [p-b] were in free variation.

Table 5.6. Mean RT (ms) and log RT on the test items (minimal pairs /t-d/ and free variants [p-b])

Test items		Orth–				Orth+			
		RT		log-RT		RT		log-RT	
L1 Group	Condition	Mean	SD	Mean	SD	Mean	SD	Mean	SD
American	Minimal Pairs	1208	443	3.06	0.13	1340	438	3.11	0.13
	Free Variants	1026	303	2.99	0.13	1413	604	3.12	0.16
Taiwanese	Minimal Pairs	2040	1263	3.24	0.23	1555	612	3.16	0.15
	Free Variants	2088	921	3.27	0.20	1755	945	3.19	0.20

Table 5.7 below shows the mean RT on minimal pairs and free variants by both American and Taiwanese participants when [t-d] were in free variation.

Table 5.7. Mean RT (ms) and log-RT on the test items (minimal pairs /p-b/ and free variants [t-d])

Test items		Orth–				Orth+			
		RT		log-RT		RT		log-RT	
L1 Group	Condition	Mean	SD	Mean	SD	Mean	SD	Mean	SD
American	Minimal Pairs	1180	463	3.04	0.15	1007	262	2.99	0.10
	Free Variants	1096	322	3.02	0.12	1030	241	3.00	0.10
Taiwanese	Minimal Pairs	1418	919	3.09	0.21	1831	935	3.21	0.22
	Free Variants	1542	1007	3.11	0.26	1843	1345	3.18	0.26

Table 5.8 below shows the mean RT on the baseline items (regardless of contrast types) by both American and Taiwanese participants in both matched and mismatched pairings and in both orthographic conditions.

Table 5.8. Mean RT (ms) and log-RT on the baseline items

Baseline items		Orth–				Orth+			
		RT		log-RT		RT		log-RT	
L1 Group	Condition	Mean	SD	Mean	SD	Mean	SD	Mean	SD
American	Matched	979	349	2.97	0.13	995	330	2.98	0.12
	Mismatched	950	267	2.96	0.11	1035	399	2.99	0.13
Taiwanese	Matched	1159	502	3.03	0.17	1279	865	3.05	0.20
	Mismatched	1152	459	3.03	0.15	1269	687	3.06	0.18

The linear mixed-effects model on the log-RT data declaring Group, OrthoCond, Item Type, Condition and Contrast Type as fixed factors and Subject and Item as random factors showed a main effect of Item Type [$F(2, 47) = 30.05$; $p < .001$], showing that the baseline items yielded the fastest RT ($M = 3.01$), minimal pairs the second fastest ($M = 3.11$) and free variants the slowest ($M = 3.12$), and of Group [$F(1, 34) = 13.88$; $p = .001$], showing that overall

Americans were faster (*M* = 3.019) than Taiwanese (*M* = 3.14). Contrast Type [$F(1, 34) = 2.70$; $p = .110$], OrthoCond [$F(1, 34) = .291$; $p = .593$] and Condition [$F(1, 1486) = 2.09$; $p = .149$] were not significant.

There were significant two-way interactions of Item Type by Contrast Type [$F(2, 42) = 3.34$; $p = .045$] where the difference of the minimal pairs and free variants between the two contrast types is much larger than that of the baseline items. Another two-way interaction of Item Type by Group [$F(2, 1545) = 13.9$; $p < .001$] indicated that Americans' RTs on the minimal pairs and free variants were about the same but Taiwanese participants were faster on the minimal pairs than free variants. There was also a significant three-way interaction of Group by OrthoCond by Contrast Type [$F(1, 33.9) = 4.37$; $p = .044$] where Taiwanese Orth– were faster than Orth+ when [t-d] were free variants but not when [p-b] were free variants. Americans, however, showed an opposite pattern: American Orth+ were faster than Orth– when [t-d] were free variants but not [p-b]. There was another significant three-way interaction of OrthoCond by Condition by Contrast Type [$F(1, 1545) = 5.86$; $p = .016$], showing that when [t-d] were free variants, the RT for the matched condition between Orth+ and Orth– were not significantly different, but when [p-b] were free variants, Orth– was faster than Orth+ on the matched condition. No other significant interactions were found.

Figure 5.4 and 5.5 below show the mean log-RT on the test items as a function of orthographic exposure by both L1 groups.

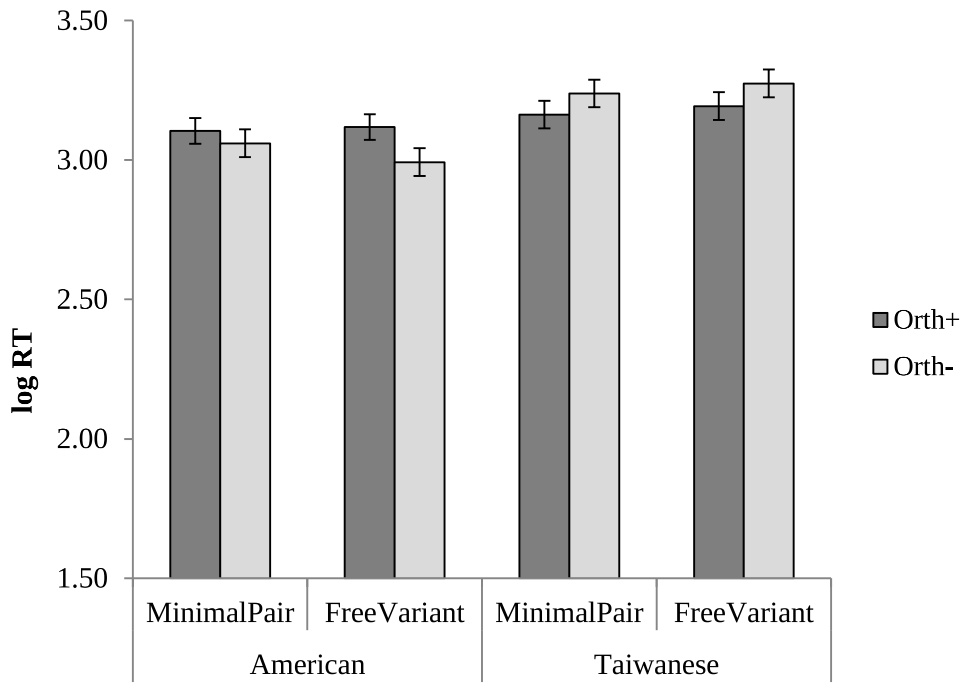


Figure 5.4. Mean log-transformed response times (log-RT) on the free variants and minimal pairs for both L1 groups as a function of orthographic exposure when [p-b] were in free variation and /t-d/ were contrastive (dark bars: Orth+; light bars: Orth–)

A post-hoc pairwise comparison with Bonferroni correction for multiple comparisons showed that Orth+ was not significantly faster than Orth– (p > .1), and American Orth– responded marginally faster than Orth+ (p = .08) on the [p-b] free variants (see Figure 5.4 above).

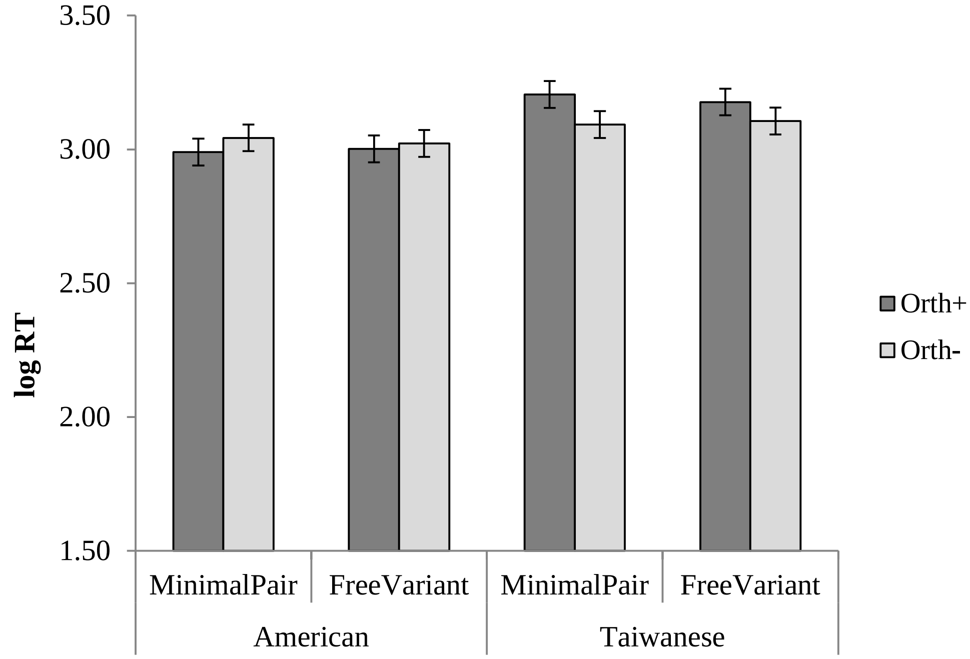


Figure 5.5. Mean log-transformed response times (log-RT) on the free variants and minimal pairs for both L1 groups as a function of orthographic exposure when [t-d] were in free variation and /p-b/ were contrastive (dark bars: Orth+; light bars: Orth–)

From Figure 5.5 above, like the previous case, Orth+ was not significantly faster than Orth–, although Taiwanese Orth– were slightly faster than Orth+ on both conditions (all $p > .1$).

Figure 5.6 below presents the log RT for the baseline items (regardless of contrast types) as a function of orthographic exposure by both L1 groups.

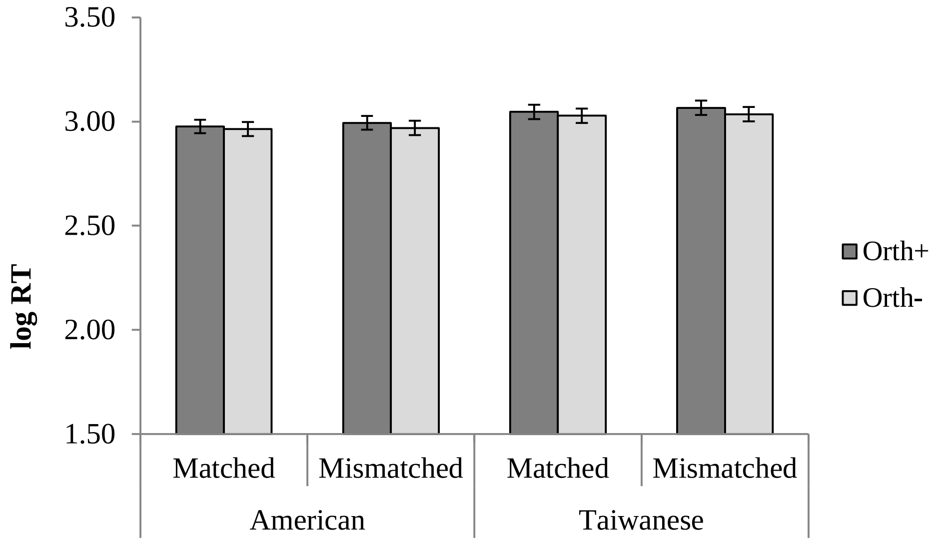


Figure 5.6. Mean log-transformed response times (log-RT) on the baseline items for both L1 groups as a function of orthographic exposure (dark bars: Orth+; light bars: Orth–)

Figure 5.6 above shows that there was no significant difference between Orth+ and Orth– in both L1 groups (all $p > .1$). In general, Americans were faster than Taiwanese participants.

In order to examine participants' sensitivity to the consonantal alternations, the next section will present the *d'* data as well as the criterion *c* to examine the potential response biases.

5.6.4 D-prime data

First we will examine the *d'* data on the test items (consonantal alternations). Figure 5.7 below shows the *d'* values on [p-b] free variation for both L1 groups in two orthographic conditions.

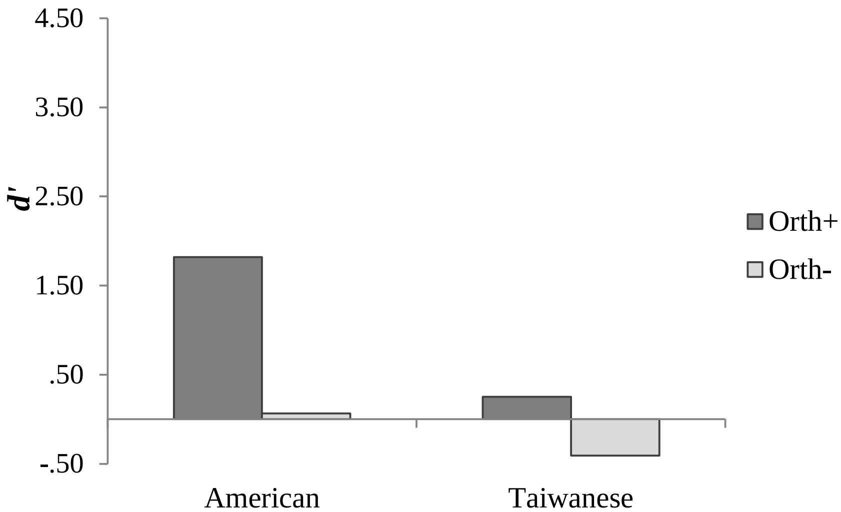


Figure 5.7. Mean *d'* for both L1 groups on consonantal alternations as a function of orthographic exposure when [p-b] was in free variation and /t-d/ in minimal pairs (dark bars: Orth+; light bars: Orth–)

As in Experiment 1, *G* test proposed by Gourevitch and Galanter (1967) was used to test the significance of the *d'* data. American participants in the Orth+ condition detected the consonantal alternations significantly better than those in the Orth– condition when [p] and [b] were in free variation (*G* = -3.98; *G* is significant at or above |1.96|). Taiwanese Orth+ participants detected the alternations slightly better than Orth– (*G* = -1.61). Notably, the American Orth+ group has a *d'* of 1.82, which is highest amongst all the groups. The result suggests that presentation of orthographic forms during learning helped American learners detect the [p-b] free variation.

Figure 5.8 below shows the *d'* values on [t-d] free variation for both L1 groups in two orthographic conditions.

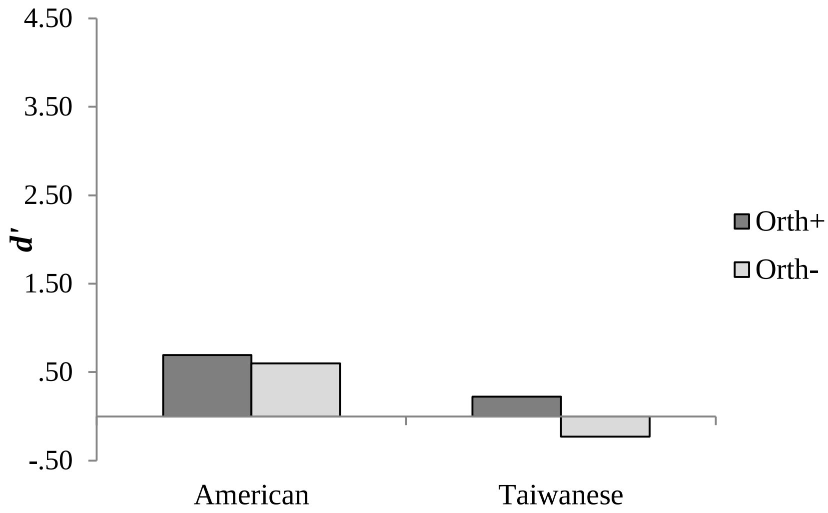


Figure 5.8. Mean *d'* for both L1 groups on consonantal alternations as a function of orthographic exposure when [t-d] was in free variation and /p-b/ in minimal pairs (dark bars: Orth+; light bars: Orth−)

Figure 5.8 above shows that when [t] and [d] were in free variation, there was no significant difference between Orth+ and Orth− in both L1 groups (both $G < |1.96|$). Also, all groups' *d'* scores were very low.

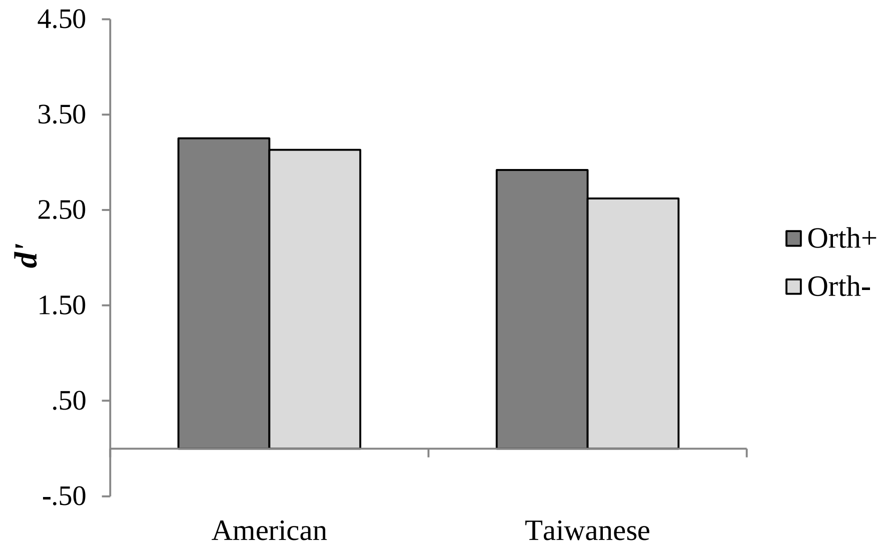


Figure 5.9. Mean *d’* for both L1 groups on baseline items as a function of orthographic exposure (dark bars: Orth+; light bars: Orth–)

From Figure 5.9 above, it can be observed that both L1 groups detected the matched and mismatched items on the baseline condition equally well regardless of their exposure to orthography, although Taiwanese Orth+ had a slightly higher *d’* than Orth–.

Next, in order to examine whether participants showed any response bias to answer either “*yes*” or “*no*”, the criterion *c* was computed. As mentioned in section 3.6.4, a negative *c* indicates “*yes*” bias while a positive *c* indicates a bias towards “*no*”.

Table 5.9 below presents the response bias, represented as criterion *c*, for both groups.

Table 5.9. Response bias in both L1 groups in both contrast types and both orthographic conditions represented as criterion *c*

	[p-b] /t-d/		[t-d] /p-b/		Baseline	
	Orth–	Orth+	Orth–	Orth+	Orth–	Orth+
American	0.221	0.613	0.299	0.409	0.072	-0.002
Taiwanese	0.471	-6.9e-17	-0.178	-0.015	-0.179	0.026

Table 5.9 above shows that Americans had some degree of "*no*" bias in both contrast types, and such bias was even stronger in Orth+ as indicated by the positive *c*, suggesting that Americans tended to split the categories. For instance, they tended to reject that [p] and [b] were free variants. In other words, they mistakenly considered that [p] and [b] were contrastive. Meanwhile, they also accepted that /t/ and /d/ were contrastive. By contrast, Taiwanese Orth+ participants have hardly any response bias in both contrast types, whereas Taiwanese Orth– participants have some degree of "*no*" bias on [p-b] free variation, suggesting a tendency to split the categories. In other words, they tended to not accept that [p] and [b] were free variants but at the same time accept that /t/ and /d/ were contrastive. Conversely, Taiwanese Orth– had some "*yes*" bias on [t-d] free variation, suggesting a tendency to merge categories. Lastly, participants had little response bias on the baseline items except for Taiwanese Orth– who had a small degree of "*yes*" bias.

In sum, the *d'* data also demonstrated an asymmetry between [p-b] and [t-d] free variation where American Orth+ showed significant better detection on the consonantal alternations than Orth– when [p-b] were free variants but not [t-d]. Taiwanese, on the other hand, showed low detection of the alternations in both contrast types, although Orth+ also showed slightly higher *d'* than Orth–. Lastly, the criterion *c* data revealed that American Orth+ showed higher degree of "*no*" bias than Orth– in both contrast types, suggesting a preference for the original forms (i.e., rejecting the new forms as the acceptable auditory labels) for the lexical items they learned. On the contrary, when Taiwanese participants were *not* exposed to orthography, they only showed the "*no*" bias on [p-b] free variation but some degree of "*yes*" bias on [t-d] free variation and baseline items, suggesting that Taiwanese Orth– preferred the original forms of [p-b] free variation but tended to accept more new forms of the [t-d] free variant as acceptable auditory labels.

5.6.5 Comparison between Experiment 1 and the current experiment

To compare whether participants learn the consonantal free variation better than the vocalic free variation, a generalized estimated equation model for the accuracy on the test items (excluding baseline) in both experiments declaring Experiment (Experiment 1 vs. 3), OrthoCond, Group, Condition and Item Type as fixed factors and Subject and Item as random factors was fitted. Note that the data of the current experiment in this analysis is not split according to the contrast type, since the main purpose of this analysis is to compare whether the accuracy on the vocalic free variation is significantly different from that on the consonantal free variation as a whole. The results show a main effect of OrthoCond [$\chi^2(1) = 6.85$; $p = .009$] where Orth+ yielded significantly higher accuracy ($M = 0.63$) than Orth– ($M = 0.52$). Lastly, Item Type was also a significant factor [$\chi^2(2) = 304.76$; $p = .001$] where the free variants yielded higher accuracy ($M = 0.70$) than minimal pairs ($M = 0.44$). There was a significant two-way interaction of Experiment by Group [$\chi^2(1) = 11.4$; $p = .019$] where in Experiment 1 (vowel alternations) Americans scored lower ($M = 0.52$) than Taiwanese participants ($M = 0.63$), but in Experiment 3 (consonant alternations) Americans scored higher ($M = 0.66$) than Taiwanese ($M = 0.49$). The interaction of Experiment by Item Type [$\chi^2(2) = 114.19$; $p < .001$][9] was also significant, in which the accuracy of the free variants was higher (*M* = 0.85) than minimal pairs (*M* = 0.25) in Experiment 1, but in Experiment 3 the accuracy of the free variants was lower (*M* = 0.49) than minimal pairs (*M* = 0.66). There was also a significant three-way interaction among Experiment, Group and Item Type [$\chi^2(2) = 20.09$; $p < .001$] in which Taiwanese participants had higher accuracy on the minimal pairs (*M* = 0.4) than Americans (*M* = 0.14) in Experiment 1, but in Experiment 3, Taiwanese participants had lower accuracy on the minimal pairs (*M* = 0.52) than Americans (*M* = 0.78).

Table 5.10 below presents the results of a comparison between Experiment 1 and 3 in terms of mean accuracy, by item types, L1 groups and orthographic conditions.

Table 5.10. Post-hoc pairwise comparison of accuracy between Experiment 1 and 3 on minimal pairs and free variants, by each L1 group and orthographic condition

OrthoCond	**Group**	**Item Type**	**Mean ACC Difference (Exp1 minus Exp3)**	***p***
Orth–	American	Minimal Pairs	-0.54	< .001
		Free Variants	0.38	< .001
	Taiwanese	Minimal Pairs	-0.08	0.403
		Free Variants	0.41	< .001
Orth+	American	Minimal Pairs	-0.71	< .001
		Free Variants	0.36	< .001
	Taiwanese	Minimal Pairs	-0.15	0.115
		Free Variants	0.29	0.001

Note: Bonferroni correction is used. A negative difference of means indicates *higher accuracy* in the *consonantal experiment*.

[9] Since the main purpose of this model is to examine the potential difference between Experiment 1 and 3, only the interactions that involve the factor "Experiment" were specified in the model.

From Table 5.10 above, it can be observed that the accuracy on the consonantal minimal pairs was higher than that on the vocalic ones across both L1 groups and orthographic conditions. Americans especially showed a highly significant increase in accuracy on the consonantal minimal pairs across both orthographic conditions. On the other hand, the accuracies on the consonantal free variants were all lower than those on the vocalic ones across both L1 groups and orthographic conditions.

The pattern of results above suggests that, compared to Experiment 3, in general participants tended to answer more "*yes*" to the vocalic minimal pairs and free variants, which results in high hit rate and false alarm rate, whereas they tended to answer more "*no*" than "*yes*" to the consonantal items, yielding high miss rate and correct rejection rate (i.e., low hit rate and low false alarm rate). To better compare the detectability of the alternations in both experiments, Table 5.11 below presents a direct comparison of the *d'* data between the vocalic and consonantal alternations.

Table 5.11. Comparison of *d'* on the vocalic (Experiment 1) and consonantal alternations (Experiment 3) across contrast types, in both orthographic conditions and both L1 groups (*G* is significant at or above |1.96|)

Group	**OrthoCond**	**Vocalic**	**Consonantal**	***G* statistic**
American	Orth–	-0.16	0.33	1.52
	Orth+	0.32	1.22	2.64*
Taiwanese	Orth–	0.63	-0.30	-3.13*
	Orth+	0.70	0.24	-1.50

Table 5.11 above shows that Americans did detect the consonantal alternations better than the vocalic ones, especially in Orth+ where the difference was significant (*G* = 2.64). Nevertheless, Taiwanese participants showed an opposite pattern: their detection of the

consonantal alternations was worse than that of the vocalic ones, especially in Orth– where the difference was significant (*G* = -3.13).

Overall, Americans showed better performance on the consonantal alternations, especially Orth+ and on [p-b] free variation. Although Taiwanese participants' detection of the consonantal alternation was lower than that of vocalic one, the orthographic effect was stronger in consonantal alternations in both L1 groups.

5.7 Discussion

To summarize, there are two major findings from the current experiment:

(1). Exposure to orthography did not always help L2 learners in the learning of the consonantal free variation.

(2). Both L1 groups did not detect the alternations very well regardless of the contrast types except for American Orth+ who showed considerable accuracy in detecting the [p-b] free variation. Taiwanese participants' detection of the consonant alternations was less accurate than their detection of the vowel alternations (as shown in Table 5.11 above).

In terms of the effect of orthography on word learning, the current study also shows mixed results. The findings from the current study provide more evidence that exposure to orthographic forms is not necessarily beneficial in word learning. As found in previous studies, some factors that determine whether orthographic information helps in word learning include L1-L2 orthography mapping (e.g., Escudero et al., 2014; Hayes-Harb et al., 2010), a possible task effect found in Simon et al. (2010), and the type of contrast as found by Escudero (2015). The current study also showed that the type of alternation (i.e., [p-b] vs. [t-d] as free variants) is a contributing factor for the effect of orthography in L2 word learning. The discussion for each research question and hypothesis is presented in order below.

5.7.1 *RQ1:* Compared to the findings in Experiment 1, can learners learn to associate the consonantal free variants with one single lexical entry more accurately better than the vocalic one?

***H1:* If consonants are tied with lexical access more than vowels as found in previous studies, it is hypothesized that learners would be able to associate the consonantal free variants more accurately than vowel alternations.**

Referring to Table 5.11, the results show that American participants did learn and detect the consonantal free variation better than the vocalic one, but not Taiwanese participants, whose detection of consonantal alternations was in fact worse than that of vocalic alternations. Thus, Hypothesis 1 is only partially confirmed.

As shown in Figure 5.1, when [p] and [b] were in free variation, American participants' accuracy on the minimal pairs was higher than that on the free variants in both orthographic conditions. Taiwanese Orth+ scored the same on the minimal pairs and free variants, but Taiwanese Orth− scored higher on the minimal pairs than free variants. These findings suggest that most of the participants in both L1 groups tended to answer "*no*" to the new variant or word paired with the picture they learned in the learning phase, an interpretation supported by the positive value of criterion *c* (except for Taiwanese Orth+) when [p] and [b] were in free variation. In other words, most of the participants tended to reject that the new variant or word they heard was a correct label for the picture when it contains [p] or [b] as the free variant.

When [t] and [d] were in free variation, the pattern of results was similar to that of [p-b] free variation except that Taiwanese Orth– scored higher on the free variants than minimal pairs, which shows a small bias toward "*yes*" responses as indicated by the negative value of criterion *c*.

Taken together, the current findings of Experiment 3 reveal that participants' recognition of the words tends to be bound by the items they learned in the learning phase, so that they tended to reject the new form of the word as the acceptable label for the lexical item they learned in the learning phase. This finding contrasts with that of Experiment 1 in which participants tended to accept the new form of the word (with vowel alternations) as an acceptable label for the lexical item they learned.

In Hypothesis 1, it was posited that learners would be able to learn the consonantal free variation better because consonants are tied with lexical access more than vowels as found in previous studies (e.g., Cutler et al., 2000). However, the current findings suggest that most learners either did not generalize what they had learned in the learning phase to the testing phase or did not detect the consonantal alternations. Instead, they tended to respond based on the original form of the word they learned in the learning phase, especially in the case of American participants. One possible reason for this finding is that in the current experiment, the alternations always occur in the word initial position, which is more perceptually salient than word medial, so learners might have a more robust representation of the initial segments than when the alternating segments are in the word medial position as in Experiment 1. Thus, they might have a robust representation of the original forms and be reluctant to accept the new forms as acceptable labels.

Alternatively, it could be that because consonants are tied with lexical access more than vowels as found in previous studies (see section 3.7.1 for more detailed discussion), learners would be able to form robust representations of the words they learned. So when the word initial segment changes, listeners would be able to detect it better and reject the new form of the original item. In this case, their exposure to the free variation rule in both the familiarization and learning phases did not significantly help them establish the link between the two variants to one lexical entry. However, one caveat is that Taiwanese only showed the preference for the original forms (i.e., a tendency to respond "*no*" to the new forms, resulting in higher accuracy on minimal pairs than free variants) in Orth– group's [p-b] free variation (see Figure 5.1 above). On the contrary, in both [p-b] and [t-d] free variation, Taiwanese Orth+ scored about 5% above chance on minimal pairs and free variants, which may suggest that they could have been influenced by the exposure to the free variation rule learned in the learning phase. If rule generalization did not come into play at all for Taiwanese Orth+, we would expect to see a pattern of results like that obtained by the Americans, where the accuracy on the minimal pairs is higher than that on the free variants.

Nevertheless, when [t] and [d] were in free variation and /p/ and /b/ were contrastive, Taiwanese Orth– scored higher on the free variants than minimal pairs, although their accuracy on the free variants was only slightly above chance level (see Figure 5.2 above). This result suggests that when Taiwanese encountered the [t-d] free variation, they were uncertain whether the new form was an acceptable label for the lexical item they learned. Nevertheless, when they encountered the new form of the minimal-pair item (i.e., /p-b/), they tended to mistakenly accept it as a correct auditory label for the word they learned, again an interpretation supported by the negative value of criterion *c* as shown in Table 5.9 above. In other words, Taiwanese Orth– tended to merge /p/ and /b/. Compared to Taiwanese Orth+ who scored above chance on both types of items, it is possible that Taiwanese participants who had not seen the orthographic forms tended to accept new phonetic form as the acceptable label for the picture they learned regardless of whether the alternation is contrastive or not, which may suggest a tendency of the benefit of orthography in the learning of free variation. However, the reason why Taiwanese tended to split [p] from [b] when [p-b] were in free variation but merge /p/ and /b/ when this pair was contrastive remains unclear given the current data.

5.7.2 *RQ2:* Will orthography benefit the learning of consonantal free variation?

> ***RQ2.1*: If the answer to Question 2 above is "yes", will there be a between-group (i.e., Taiwanese vs. American) difference in the effect of orthography and the learning of free variation?**

***H2:* Because the grapheme-to-phoneme correspondence of consonants in English is more transparent than that of vowels, learners would benefit more from the orthographic information in detecting the consonantal alternations and learning the free variation.**

> ***H2.1*: Since the target language's grapheme-to-phoneme correspondence of the consonants matches with that of English, American participants would show a larger effect of orthography than Taiwanese participants.**

The results show that when [p] and [b] were in free variation, exposure to orthographic information during the familiarization and learning phases significantly helped participants learn that [p] and [b] were free variants while /t/ and /d/ were contrastive. Also, the effect of orthography is particularly stronger for American compared to Taiwanese participants. Although Taiwanese participants could use the orthographic information to help them learn that [p] and [b] were free variants, it did not help them learn that /t/ and /d/ were contrastive. In fact, although not significant, Taiwanese participants without exposure to orthography scored slightly higher than those who were exposed to orthography on the minimal pair /t/ and /d/. In general, *d'* scores showed that orthographic information did benefit both groups, especially Americans, in detecting the [p-b] free variation.

However, when the free variants were [t] and [d] and the minimal pair was /p/ and /b/, orthographic information did not significantly help both groups in detecting the alternation, although it slightly helped Taiwanese participants learn that /p/ and /b/ were contrastive.

Taken together, the current findings showed mixed results as to the effect of orthography on the learning of the consonantal free variation. One possible account for this asymmetry between the two types of contrast (labial vs. coronal) is the frequency of occurrences of the coronals. According to Paradis and Prunet (1991) and Keating (1991), the frequency of occurrences of coronals is higher than bilabials cross-linguistically. In other words, coronals are typologically more frequent than bilabials. In addition, coronals are more subject to phonological variations (e.g., assimilation) than consonants of other places of articulation (Stemberger & Stoel-Gammon, 1991). Thus, when learners encounter the free variants /p/ and /b/, which are relatively less likely to undergo phonological variation than coronals /t/ and /d/, orthography may help them link the variants containing these two sounds to one single lexical entry. On the contrary, when they encounter a less marked pair of variants (i.e., /t/ and /d/), the learning outcome may not be affected by the exposure to the orthographic information during the learning process. However, one caveat is that the phonological alternations of coronals occur mostly in word- or sentence-medial or word-final positions, so it is questionable whether similar phenomena would generalize to word-initial position as in the invented language in the current experiment. The second caveat is that both L1 groups still did not detect the [t-d] free variation very well. If coronals are less marked than bilabials, presumably participants should be able to learn the [t-d] free variation better than [p-b], which is not the case in the current study, except for American Orth− who detected the [t-d] free variation slightly better than [p-b], which could be due to the English flapping rule (discussed below).

Another possible account for the asymmetry between the two types of contrast is that Americans are familiar with the English flapping rule in which both /t/ and /d/ can be realized as [ɾ] when it occurs inter-vocalically and before an unstressed vowel (Ladefoged, 2000). When orthographic forms were not presented, Americans scored 60% on /t-d/ minimal pairs but 73% on /p-b/ minimal pairs. In other words, when no orthography is available, Americans are more likely to split /p/ from /b/ than to split /t/ from /d/, which may be due to the L1 influence from the flapping rule. Taiwanese Orth–, however, scored 61% on /t-d/ minimal pairs but 38% on /p-b/ minimal pairs, suggesting that they are more likely to split /t/ from /d/ than /p/ and /b/, showing an opposite pattern to that of American participants. This may suggest that the English flapping does not have a significant impact on Taiwanese participants' responses.

Lastly, given limited funding as well as time constraints, the sample size in the current experiment is relatively small. The original design of this experiment was to counterbalance the contrast types and collapse the data from the two contrast types for analysis. Nevertheless, a main effect of contrast type was later found in the analysis, so the results of the two contrast types were interpreted separately. The separation of the results in turn reduces the sample size of each subgroup (for each contrast type, there were 5 people from each L1 group in each orthographic condition), which also reduced statistical power. In future research, more participants are needed to reach higher predictive power.

Chapter 6. Experiment 4; Production of words with consonant alternations

As in Experiment 2, the goal of the current experiment is to investigate whether exposure to orthographic forms helps participants spontaneously produce the words they learned in the learning phase and produce new forms of the free variants.

The specific research questions to be answered in this experiment are:

1. Can learners who have exposure to orthographic information be more accurate in picture naming than learners who do not?
2. Would exposure to orthography cause any differences in the encoding the free variation? In other words, will exposure to orthography help learners produce more new forms of the free variant?

The hypotheses are:

1. Exposure to orthographic information will help learners encode and decode the words better.
2. Exposure to orthography will help learners produce more new forms of the free variant as well as produce the original form of the minimal-pair item.

In addition to the research questions above, the results from the current experiment will be compared to that of Experiment 2 to examine any potential differences in picture naming involving vowel and consonant alternations.

6.1 Materials and Design

This experiment parallels Experiment 2 in terms of design. Materials for the familiarization and learning phases were the same as those used in Experiment 3. However, in the picture naming task, for the items of minimal pairs, only the pictures from the learning phase were presented, since participants only learned one word from the pair (and the other word – the new token – was used to create a mismatched trial.) Each item appeared twice, yielding a total of 32 trials. There were 8 trials of minimal pair items, 8 trials of free variants and 16 baseline trials.

6.2 Participants

Participants were the same as in Experiment 3 (see Chapter 5), and the exclusion criteria from Experiment 3 also apply in the current experiment.

6.3 Procedure

Procedure was the same as in Experiment 2 (see Chapter 4).

6.4 Analysis

The scoring scheme and the methods of analysis were similar to that in Experiment 2. There were a total of 1312 responses recorded. No response was discarded. The coding for consonant production was the same as in Experiment 2. However, for word-initial consonants where the experimental segments occurred, unlike in word-final positions, a change from [t, d] to [p, b] was not accepted, and vice versa, because the two pairs of consonants (i.e., [p, b] and [t, d]) were in two different experimental conditions and should not be an acceptable deviation. As mentioned in Chapter 4, the data of naming accuracy (i.e., production score) provides different information than the data of the production of the target alternation (i.e., the frequency of occurrences of the target alternation production). So, for example, for [t-d] free variation, if the free variant is [tisɛf] and participant produces [dumak], production of new form of free variant ([d]) is coded for this response even though the rest of the sounds are not correct.

The recordings were first transcribed by the author with the aid of spectrograms. In specific cases where the word-initial stops could not be clearly judged as either aspirated or unaspirated, that trial would be coded separately for later analysis. Such trials (6 out of 1312 total trials) were included in the analysis of naming accuracy. However, these 6 trials were dropped from the analysis of target alternation production, considering that the unclear aspiration would confound the scoring of the free-variant production. For example, if one learned [p] as a free variant and produced [p] with a shorter aspiration, which may make it sound like [b], it would be difficult to judge whether s/he produced an original or new form of the free variant.

As in Experiment 2, a phonetically trained transcriber (the same person as in Experiment 2) transcribed four participants' recordings (out of 41). His transcriptions were then compared to mine as a reliability measure. It showed that our transcriptions reached 95% agreement.

6.5 Results

In both of the section 6.5.1 and 6.5.3, the data of the observed means and standard deviations of the production score and naming latency will be presented in tables followed by bar graphs where the means are estimated marginal means from the statistical models. In section 6.5.2, the observed proportion of each response type will be first presented in a table then in a bar graph.

6.5.1 Production score data

Table 6.1 below shows the mean production score (shown as the average number of correctly produced segments in each word) by both L1 groups in two orthographic conditions and three item types. In addition, the upper half of the table shows the data of [p-b] free variation whether the lower half shows the data of [t-d] free variation.

Table 6.1. Mean production scores in two contrast types by both Groups and orthographic conditions (perfect score = 5)

Contrast type	Groups	Item Type	Orth−	Standard deviation	Orth+	Standard deviation
[p-b] free variation; /t-d/ contrast	American	Free Variant	2.66	1.75	2.58	2.04
		Minimal Pair	2.73	1.83	3.53	1.69
		Baseline	2.32	1.92	3.10	2.04
	Taiwanese	Free Variant	3.03	1.80	2.50	2.19
		Minimal Pair	3.43	1.38	3.25	1.97
		Baseline	3.06	1.76	2.63	1.96
[t-d] free variation; /p-b/ contrast	American	Free Variant	3.70	1.67	3.00	1.95
		Minimal Pair	3.29	1.71	2.41	1.94
		Baseline	3.49	1.68	3.16	1.82
	Taiwanese	Free Variant	4.09	1.08	2.13	1.92
		Minimal Pair	2.51	1.72	3.06	1.80
		Baseline	2.46	1.70	3.00	1.80

A linear mixed-effects model was fitted to examine the effects of the fixed factors Group, OrthoCond, Item Type and Contrast Type and random factors Subjects and Items on production score (naming accuracy). The results showed that there was no main effect of Group [$F(1, 34) = .068$; $p = .795$], of OrthoCond [$F(1, 34) = .453$; $p = .506$], of Item Type [$F(1, 28) = .085$; $p = .919$] or of Contrast Type [$F(1, 34) = .195$; $p = .662$]. A significant two-way interaction of OrthoCond by Item Type [$F(2, 1242) = 11.08$; $p < .001$] was found. When Contrast Type and L1 groups were held constant, Orth– scored higher than Orth+ on the free-variant but not on the minimal-pair and baseline items. In addition, a significant four-way interaction of OrthoCond by

Group by Item Type by Contrast Type was found [$F(7, 355) = 2.99$; $p = .005$], where American Orth– scored higher than Orth+ on both minimal-pair and free-variant items when [t-d] were free variants but not when [p-b] were free variants. However, Taiwanese Orth+ scored higher than Orth– on minimal-pair items when [t-d] were free variants but not when [p-b] were free variants.

There was also a marginal three-way interaction of OrthoCond by Group by Item Type [$F(2, 1242) = 2.56$; $p = .078$].

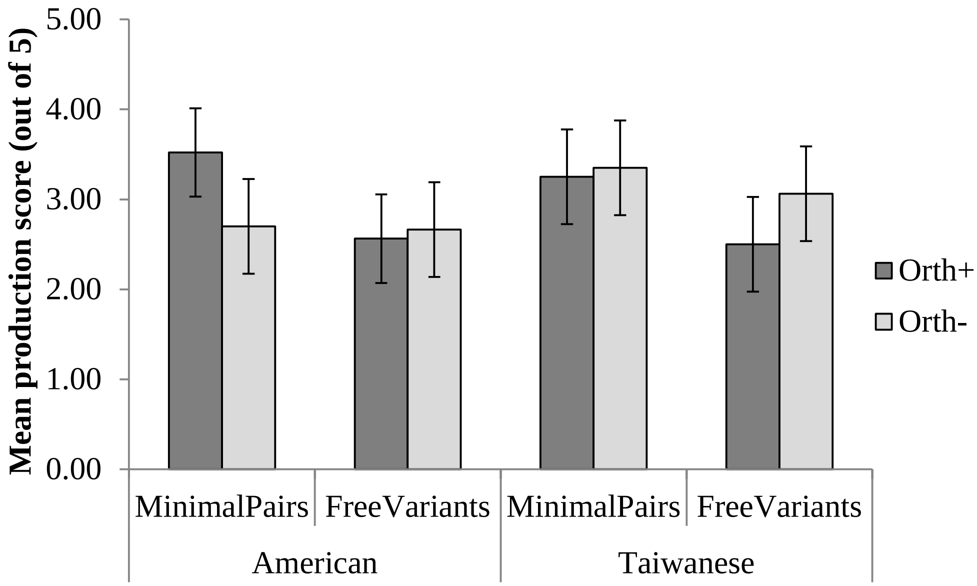


Figure 6.1. Mean production score for the test items ([p-b] free variation and /t-d/ contrastive) by both L1 groups (dark bars: Orth+; light bars: Orth–; error bars represent the standard errors of the estimated marginal means)

Although there was no main effect of orthography, Figure 6.1 above shows that when [p-b] were in free variation and /t-d/ were contrastive, American Orth+ tended to score higher than Orth− on minimal-pair items, but Taiwanese Orth− participants tended to score higher than Orth+ on free-variant items.

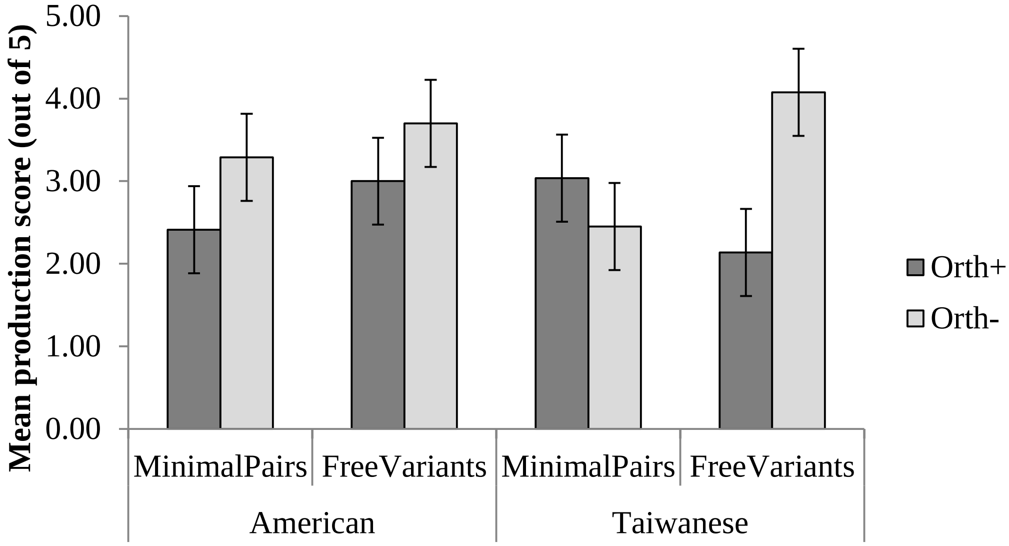


Figure 6.2. Mean production score for the test items ([t-d] free variation and /p-b/ contrastive) (dark bars: Orth+; light bars: Orth–; error bars represent the standard errors of the estimated marginal means)

When [t-d] were in free variation, a post-hoc pairwise comparison using Bonferroni correction for multiple comparisons shows that Taiwanese Orth− scored significantly higher than Orth+ on free-variant items (p = .004), suggesting that exposure to orthography hinders the production of the free-variant items. As demonstrated in Figure 6.2 above, the same pattern is also shown in Americans' production scores in which Orth− tended to score higher than Orth+. The only exception is the Taiwanese participants' production score of minimal-pair items where Orth+ scored slightly higher than Orth−.

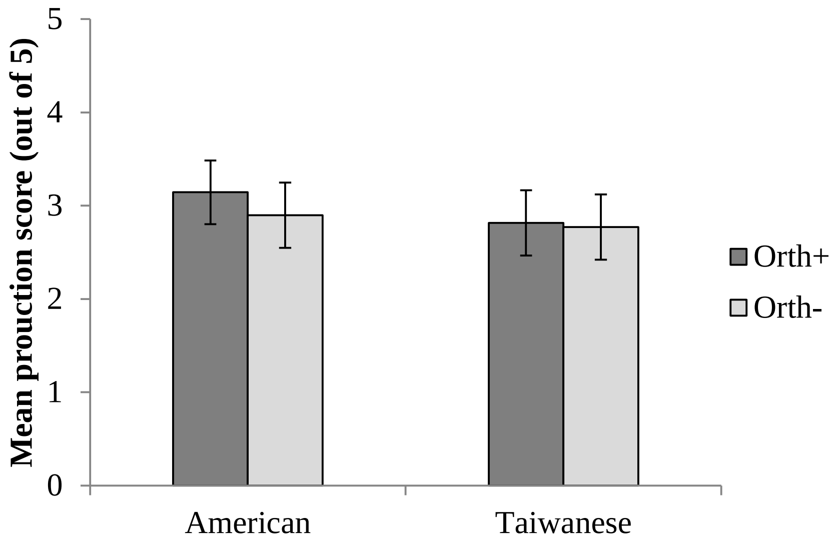


Figure 6.3. Mean production score for the baseline items by both L1 groups across both contrast types (dark bars: Orth+; light bars: Orth–; error bars represent the standard errors of the estimated marginal means)

Figure 6.3 above shows that there is no significant difference between Orth+ and Orth− in both L1 groups on the baseline items (all $p > .05$).

6.5.2 Production of the consonant alternations

The production score data shown above does not reflect participants' actual production accuracy on the alternating sounds of the two contrast types (i.e., /t-d/ or /p-b/). Thus, in order to examine what forms of the alternating sounds participants produced, as was done in Experiment 2, their responses on the alternating sounds were coded as "wrong minimal pair" in which they produced a wrong consonant in the minimal pair (e.g., learned [t] in the learning phase but produced [d], when /t-d/ are contrastive), or "original forms" in which they produced the forms they learned in the learning phase, or "new forms of the free variants" (e.g., learned [p] but produced [b] for free variants). Table 6.2 below shows the proportion of these three types of responses by both L1 groups in two orthographic conditions and in two contrast types.

Table 6.2. Proportion of three types of responses on the target consonantal alternation

Contrast type	Group	Orthographic condition	Wrong minimal pair	Original forms	New form of the free variants
[p-b] free variation; /t-d/ contrast	American	Orth+	0.01	0.94	0.04
		Orth–	0.03	0.93	0.03
	Taiwanese	Orth+	0.00	1.00	0.00
		Orth–	0.08	0.83	0.09
[t-d] free variation; /p-b/ contrast	American	Orth+	0.00	0.96	0.04
		Orth–	0.00	0.98	0.02
	Taiwanese	Orth+	0.02	0.82	0.16
		Orth–	0.06	0.94	0.00

For the analysis on the proportion of these three types of responses, the data distribution is multinomial, so a generalized estimated equation model for the three types of responses on the target alternations declaring Group, OrthoCond and Contrast Type as fixed factors and Subject and Item as random factors was fitted. The results showed that Group [$\chi^2(1) = .038$; $p = .845$], OrthoCond [$\chi^2(1) = 2.02$; $p = .155$] and Contrast Type [$\chi^2(1) = .212$; $p = .645$] were not significant. No significant interactions were found, although OrthoCond by Group by Contrast Type was marginally significant [$\chi^2(1) = 3.023$; $p = .082$].

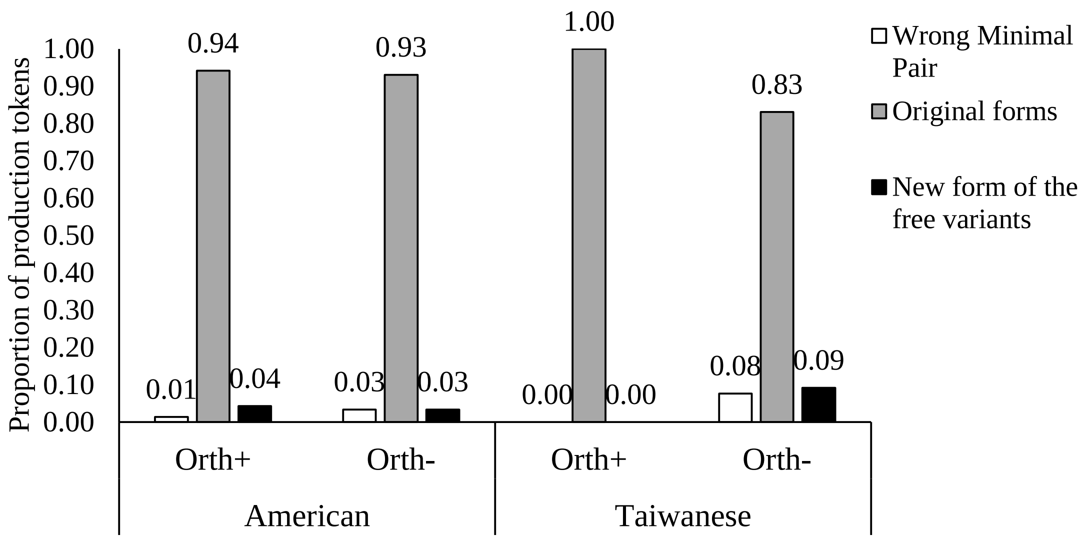


Figure 6.4. Proportion of three types of responses on the target alternation when [p-b] were free variants and /t-d/ were contrastive by both L1 groups and both orthographic conditions (black bars: production of new form of the free variant; grey bars: production of the original forms of both free-variant and minimal-pair items; white bars: production of the wrong consonant in the minimal pair)

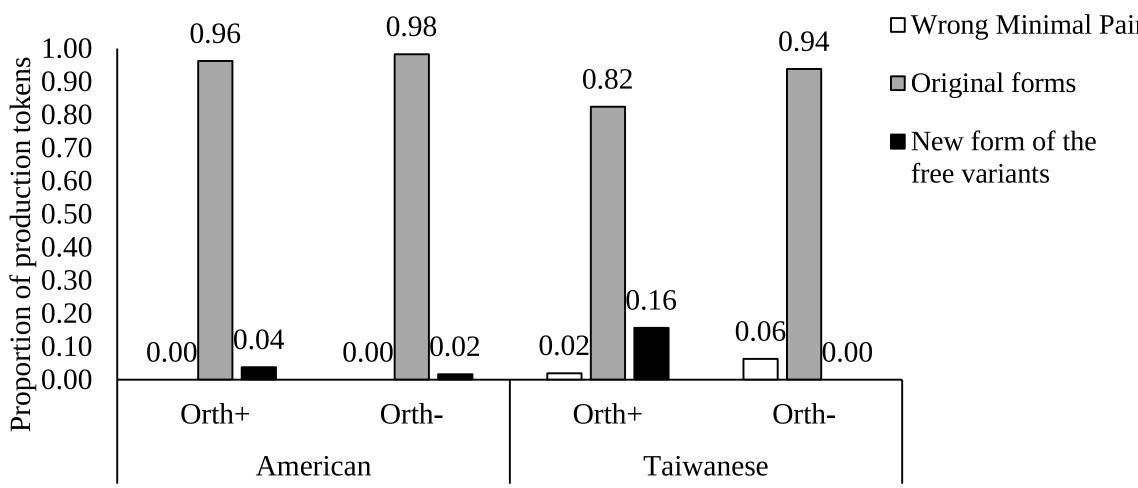


Figure 6.5. Proportion of three types of responses on the target alternation when [t-d] were free variants and /p-b/ were contrastive

From both Figure 6.4 and 6.5 above, it can be seen that the original forms that participants learned is the most frequent type of responses on the target alternations. There were only a small number of the productions of the wrong minimal pair and new form of the free variants. When [t-d] were in free variation, Taiwanese Orth+ produced more new forms of the free variants as well as fewer wrong minimal pairs than Orth−, which is the only observable benefit of orthographic information.

6.5.3 RT data

For the RT data analysis, the trials where participants did not make a response were dropped from the analysis. The distribution of the raw RT data was positively skewed, so RT data was logarithmically transformed before fitting the linear mixed-effects model. The distribution of the log-transformed data was close to normal.

Table 6.3. Mean RT (ms) and log-RT by both groups and orthographic conditions

			Orth−				Orth+			
			RT		log-RT		RT		log-RT	
Contrast type	Groups	Item Type	Mean	SD	Mean	SD	Mean	SD	Mean	SD
[p-b] free variation; /t-d/ contrast	American	Free Variant	1964	883	3.26	0.18	2173	1110	3.29	0.21
		Minimal Pair	2281	1428	3.28	0.25	1848	933	3.22	0.20
		Baseline	1650	935	3.17	0.20	1917	953	3.24	0.20
	Taiwanese	Free Variant	2022	917	3.26	0.19	2124	1004	3.27	0.24
		Minimal Pair	2124	1149	3.27	0.22	2369	1117	3.33	0.21
		Baseline	2119	1041	3.28	0.20	2425	1330	3.32	0.23
[t-d] free variation; /p-b/ contrast	American	Free Variant	1523	823	3.13	0.20	1655	886	3.17	0.20
		Minimal Pair	1720	879	3.18	0.23	1623	753	3.17	0.19
		Baseline	1775	1055	3.18	0.24	1740	1052	3.18	0.22
	Taiwanese	Free Variant	1658	865	3.17	0.19	2549	1482	3.34	0.24
		Minimal Pair	1976	1001	3.24	0.22	2086	1139	3.26	0.23
		Baseline	1943	929	3.24	0.20	2419	1290	3.32	0.23

Table 6.3 above shows both the raw and log-transformed mean RT and standard deviations in both L1 groups, orthographic conditions as well as the two contrast types.

A linear mixed-effects model was fitted to examine the effects of the fixed factors Group, OrthoCond, Item Type and Contrast and random factors Subjects and Items on naming latency.

The results revealed a significant main effect of Group [$F(1, 34) = .4.45$; $p = .042$] where Taiwanese were slower than American, but there were no main effects of OrthoCond [$F(1, 34) = 1.661$; $p = .206$], Item Type [$F(1, 49) = .027$; $p = .974$] or Contrast Type [$F(1, 34) = 2.532$; $p = .121$]. No significant interactions were found, but OrthoCond by Item Type was marginally significant [$F(1, 976) = 2.587$; $p = .076$].

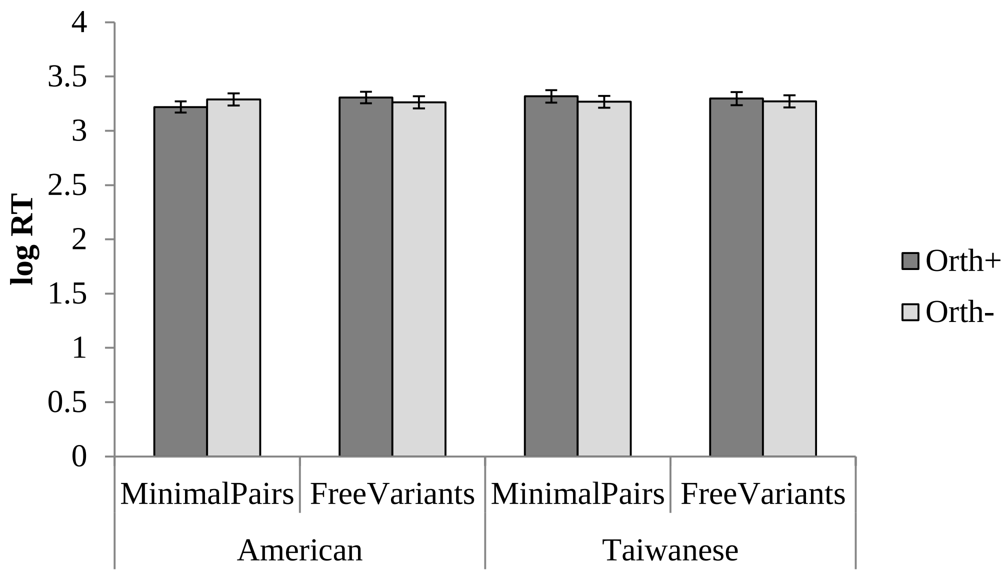


Figure 6.6. Mean naming latency (log-RT) for the test items ([p-b] free variation and /t-d/ contrastive) (dark bars: Orth+; light bars: Orth–; error bars represent the standard errors of the estimated marginal means)

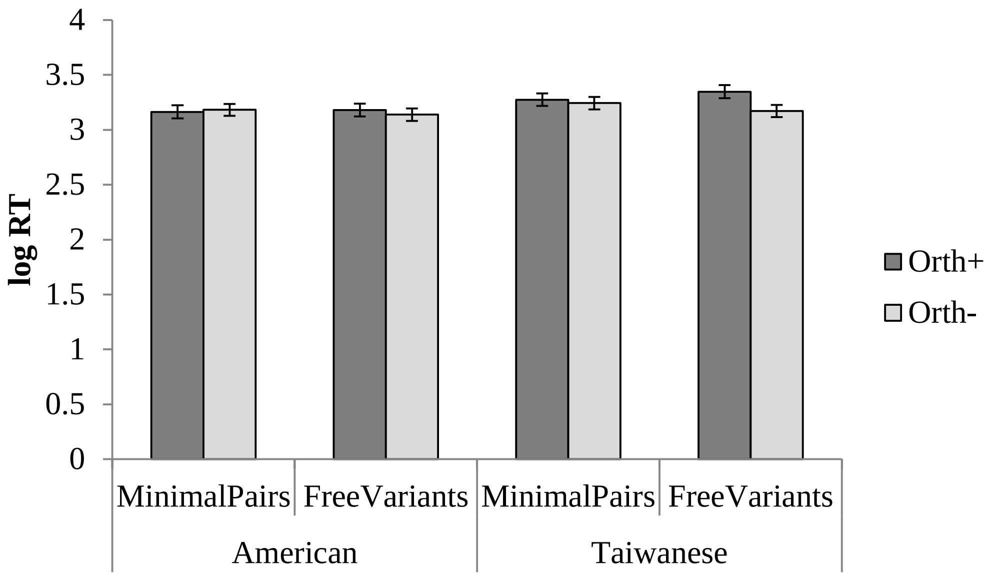


Figure 6.7. Mean naming latency (log-RT) for the test items ([t-d] free variation and /p-b/ contrastive)

A post-hoc pairwise comparison using Bonferroni correction showed that when [t-d] were in free variation Taiwanese Orth− responded significantly faster than Orth+ on free-variant items (p = .029), suggesting that exposure to orthography hinders the processing of lexical retrieval (see Figure 6.7 above).

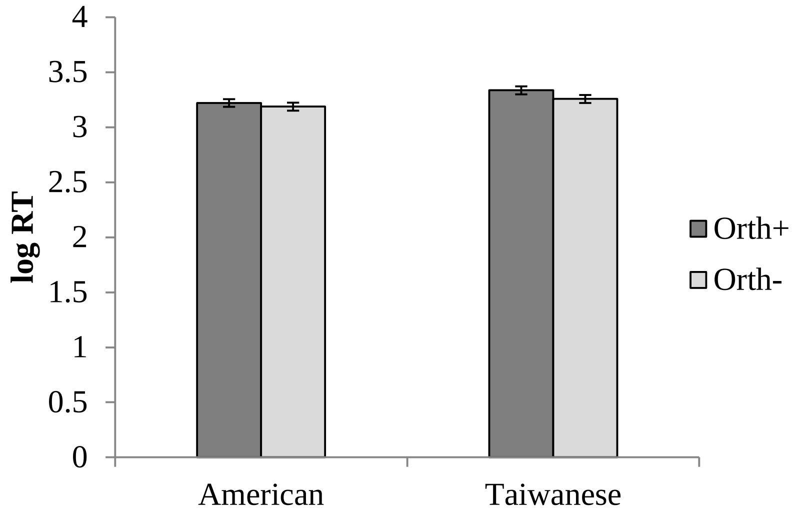


Figure 6.8. Mean naming latency (log-RT) for the baseline items and both L1 groups (dark bars: Orth+; light bars: Orth–; error bars represent the standard errors of the estimated marginal means)

Figure 6.8 above shows that there was no significant difference in naming latency for the baseline items between the two orthographic conditions in both L1 groups, suggesting that exposure to orthography did not facilitate the lexical retrieval speed.

6.5.4 Comparison between picture-word matching (Exp. 3) and picture naming (Exp. 4)

As shown in Chapter 4 (Experiment 2), in order to compare the results between Experiment 3 and 4, two scatter plots that plot each participant's overall accuracy (across all item types) on the picture-word matching (Experiment 3) against the overall accuracy on picture naming (Experiment 4) for the two contrast types are shown below.

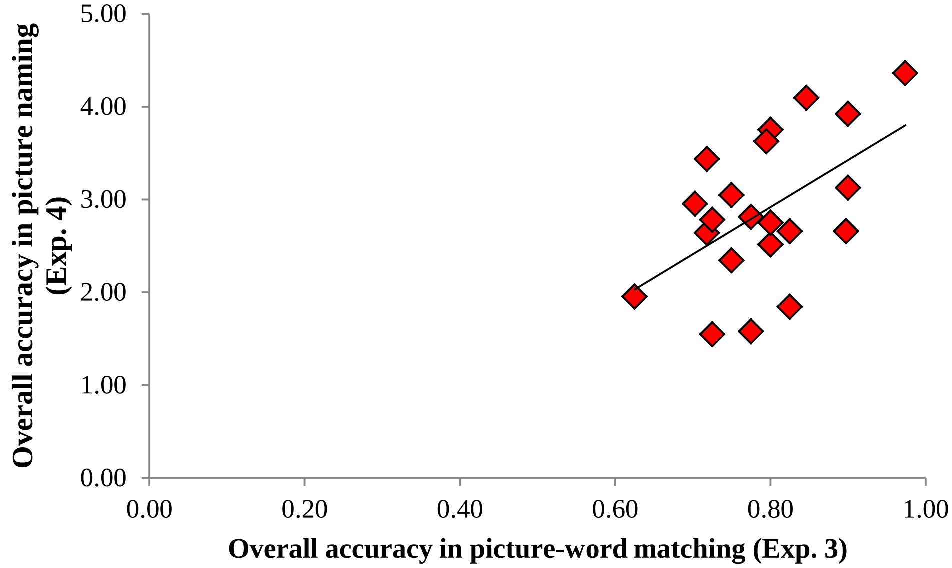


Figure 6.9. Correlation between the overall accuracy in picture-word matching and picture naming when [p-b] were in free variation and /t-d/ were contrastive; each diamond represents each participant's overall accuracy

The Pearson's *r* shows that there was a significant correlation (r = 0.507; two-tailed p = .019) between the overall accuracy in picture-word matching and picture naming when [p-b] were in free variation and /t-d/ were contrastive, suggesting that those who perform better in the picture-word matching task also scored higher in the picture naming (see Figure 6.9 above).

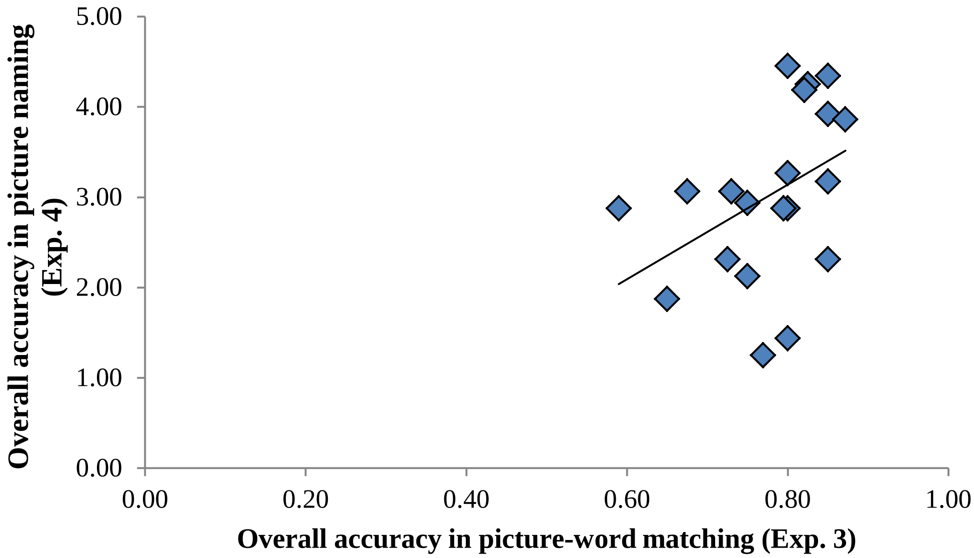


Figure 6.10. Correlation between the overall accuracy in picture-word matching and picture naming when [t-d] were in free variation and /p-b/ were contrastive; each diamond represents each participant's overall accuracy

The Pearson's *r* shows that there was a marginally significant correlation ($r = 0.414$; two-tailed $p = .07$) between the overall accuracy in picture-word matching and picture naming when [t-d] were in free variation and /p-b/ were contrastive (see Figure 6.10 above).

6.6 Discussion

The current experiment tested participants' naming accuracy of the pictures that they learned from the learning phase in Experiment 3. In addition, one important goal was to examine participants' production of the consonantal alternations to see how well they encoded and decoded the alternations. In this section, I will first discuss the general results and then the specific research questions. Finally, in section 6.6.3, participants' performance in Experiment 3 (the picture-word matching task) and the current experiment will be compared by discussing the correlation between the accuracy in Experiment 3 and the current experiment.

The results showed that when [t-d] were in free variation, exposure to orthography seemed to hinder learners' lexical retrieval and production of the words that involved free variations. A similar pattern can also be observed when [p-b] was in free variation: Orth– scored higher than Orth+ on the [p-b] free variants, especially for Taiwanese participants. The inhibitory effect of orthography on the production of the items containing free variants seems to be consistent with the finding in Damian and Bowers (2003)'s first experiment. They found that if one sound was spelled in two different ways (e.g., [k] spelled as either <c> or <k>), inconsistent spelling did not show priming effect in a form-preparation paradigm compared to the condition where both spelling and sound matched (e.g., [k] spelled only as <k>). However, there were two caveats regarding the comparison between the current findings and the Damian and Bower's study. First, if it is the inconsistent grapheme-to-phoneme correspondence that causes the inhibition effect, a similar pattern of results (i.e., Orth– scored higher than Orth+) should *not* show up in the results of minimal pairs because in minimal pairs, the grapheme-to-phoneme correspondence for the alternating sounds is consistent. Nonetheless, American Orth– also scored higher than Orth+ on the items of /p-b/ minimal pairs (i.e., when [t-d] were in free variation) (see Figure 6.2), suggesting an inhibitory effect of orthography, which is unexpected since [p] and [b] were spelled as <p> and <b> respectively when /p-b/ were contrastive. Further discussion on this observation is presented under Research Question 1 below.

The second caveat is the production of the alternating sounds. Figure 6.5 shows that when [t-d] were in free variation, both Americans and Taiwanese Orth+ produced more new forms of the free variants than Orth– (although the proportion of such type of response is relatively low), which seems to be inconsistent with the previous finding that inconsistent spelling will not help in word production. Some possible explanations are discussed under Research Question 2 below.

6.6.1 *RQ1*: Can learners who have exposure to orthographic information be more accurate in picture naming than learners who do not?

***H1*: Exposure to the orthographic information will help learners encode and decode the words better.**

The current results did not show a clear benefit of orthographic exposure on picture naming. On the contrary, when [t-d] were in free variation, exposure to orthography in fact significantly hindered Taiwanese participants' production accuracy on the free-variant items. A similar trend was also found in Americans' production where the Orth+ group was overall less accurate on both free-variant and minimal-pair items when [t-d] were in free variation, compared to when [p-b] were in free variation. As shown in Figure 6.1, the only benefit of orthography was visible in Americans' production of /t-d/ minimal pairs (i.e., when [p-b] were in free variation) and Taiwanese' production of /p-b/ minimal pairs (i.e., when [t-d] were in free variation). Thus, the first hypothesis is only partially confirmed.

Recall that Experiment 2 (i.e., picture naming on the vocalic free-variation items; see Chapter 4) revealed a benefit of orthography in word production for both L1 groups and all three types of items (i.e., free variants, minimal pairs and baseline items), but such a benefit is not prevalent in the current results. As mentioned above, the benefit of orthography in the current experiment can only be seen in Americans' production of /t-d/ minimal pairs and Taiwanese' production of /p-b/ minimal pairs. Again, if inconsistent grapheme-to-phoneme correspondences cause inhibition in picture naming as shown in Damian and Bowers' study, Experiment 2 should also show such an effect (and especially for free variants). One possible account for the asymmetric results between Experiment 2 and 4 (the current experiment) might be the nature of L1 orthographic depth. As mentioned before, Erdener and Burnham (2005) found that an inhibitory effect of orthography on production arises when L1 and L2 orthographic depth mismatches, especially when L1 orthography is transparent and L2 orthography is opaque. In Experiment 2, the alternating sounds were vowels, whose grapheme-to-phoneme correspondence was more inconsistent than consonants in English. In addition, both groups of participants in the current study were familiar with the English orthography, so they might be also more familiar with the opaque grapheme-to-phoneme correspondence of English vowels as well as the transparent orthography of consonant in English. Thus, when participants in Experiments 1 and 2 learned the novel words, they might have relied on their experience with the inconsistent grapheme-to-phoneme correspondence of vowels (i.e., <o> corresponds to [ɔ] and [u]), so that orthographic information had facilitative effect in production. On the other hand, when participants in Experiments 3 and 4 encountered the inconsistent orthography of consonantal free variants in the current experiment (i.e., <t> corresponds to both [t] and [d] when [t-d] were free variants, or <p> corresponds to both [p] and [b] when [p-b] were free variants), the presentation of the orthographic forms might have hindered their overall production accuracy. However, when the grapheme-to-phoneme correspondence is consistent in the minimal pairs, the presentation of the orthographic forms may help (e.g., the production of /t-d/ minimal pairs by Americans and the production of /p-b/ minimal pairs by Taiwanese; see Figure 6.1 and 6.2). Nevertheless, note that the benefit of orthography was not visible for both groups in both types of minimal pairs, the reason accounting for which remains unclear given the current data available.

6.6.2 *RQ2:* Would exposure to orthography cause any differences in the encoding of free variants? In other words, will exposure to orthography help learners produce more new forms of the free variants?

***H2:* Exposure to orthography will help learners produce more new forms of the free variants as well as produce the original form of minimal-pair items.**

Regardless of the contrast types (/t-d/ or /p-b/) and orthographic conditions, both L1 groups overwhelmingly produced the original forms of the test items that they learned in the learning phase as shown in Figure 6.4 and 6.5 above. The proportion of production of the new forms of free variants as well as the wrong minimal pairs was very low compared to that for the original forms. However, although not significant, when [t-d] were in free variation, Taiwanese participants who were exposed to orthography produced more new forms of free variants and fewer wrong alternating consonants in the minimal pairs than those who were not exposed to orthography, suggesting that orthographic forms might help Taiwanese participants encode and produce the free variants.

Nevertheless, when [p-b] were in free variation, Taiwanese participants who were not exposed to orthography produced both more new forms of the free variants (9%) and also more wrong alternating consonants in the minimal pairs (8%) than those who were exposed to orthography. On the contrary, Taiwanese Orth+ in this case only produced the original forms of both minimal pairs and free variants. Hence, there is neither facilitative nor inhibitory effect of orthography on Taiwanese participants' production of the alternating consonants when [p-b] were in free variation.

Taking the results from both contrast types together, the second hypothesis can only be partially confirmed. One might wonder why orthographic information still seems to help participants produce more new forms of free variants, since the inconsistent grapheme-to-phoneme correspondence of consonants in the current study might account for the low production accuracy of the free-variant items as discussed earlier. As mentioned in section 6.4, the production of the alternating sounds only shows the frequency of occurrences of producing the alternating sounds rather than the accuracy of the production. Looking at the results of production accuracy and the production of alternating sounds, we might conclude that, when L1 and L2 orthographic depth mismatches (the consonants in this case), the inconsistent grapheme-to-phoneme correspondence of the free variants might have a negative effect on the overall quality of storage for the whole words containing free variants (as shown by the lower accuracy on free-variant items by Orth+ than Orth–), but meanwhile it also slightly increases the likelihood of producing the new forms of the free variants and reduces the likelihood to produce a wrong consonant in the minimal pairs. In other words, the current results show that although inconsistent grapheme-to-phoneme correspondence can hinder picture naming as demonstrated in previous studies (especially when L1 and L2 orthographic depth mismatches), learners also seem to encode the alternations better if they were exposed to orthographic forms.

6.6.3 Correlation between recognition (picture-word matching) and production (picture naming)

The current results show that orthographic information did not always help participants correctly name the pictures. But in Experiment 2 (naming of pictures containing vowel alternations), orthographic information was found to be helpful in participants' production of vowel alternations and words. In addition, there was no correlation between the overall accuracy in the picture-word matching and picture naming, suggesting that better performance in the picture-word matching task does not imply better performance in picture naming. Furthermore, as reflected by the main effect of orthographic condition in Experiment 2, participants who were exposed to the orthographic forms were able to produce more correct forms of the minimal pairs as well as more new forms of the free variants than Orth–. In general, the findings from Experiment 2 show that orthographic information plays a role in lexical retrieval and encoding, but this effect is not visible in Experiment 1 (picture-word matching). Nevertheless, such a correlation between production and recognition was found in the current experiment as shown in Figures 6.9 and 6.10 above, and the correlation was strongest when [p-b] were in free variation. In other words, better performance in picture-word matching could predict better performance in picture naming for items containing consonantal alternations but not vowel alternations. The exact nature of this difference in correlation has yet to be further investigated. One possible explanation would be the different roles of consonants and vowels in lexical access where consonants were shown to be tied with lexical access more than vowels (e.g., Bonatti et al., 2005; Cutler et al., 2000). In the current design, the picture naming task was always preceded by the picture-word matching, so it is likely that completing the picture-word matching involving consonant alternations would also help participants remember and produce words in the picture naming task. Thus, those who performed better in picture-word matching also performed better in picture naming for words containing consonant alternations; but the picture-word matching involving vowel alternations did not help in production because vowels carry lesser weight than consonants in lexical access. One possible future direction would be to reverse the order of picture-word matching and picture naming in order to substantiate this explanation. If this account is valid, then a correlation between the two tasks involving consonant alternations should still be found even if the order is reversed, because in that case, picture naming would also help participants in picture-word matching.

Chapter 7. General Discussion and Conclusion

7.1 Summary

The current study set out to investigate: 1) the role of orthographic information in lexical access and speech perception; 2) whether orthographic information can help L2 learners (Taiwanese and American participants) establish a single lexical representation for phonetic variants in an artificial language; and 3) whether it can help L2 learners better learn the words in a second language. To my knowledge, this study is the first that investigates whether exposure to L2 orthography can help L2 learners link two phonetic variants to one lexical entry, and also one of the few studies that examines both recognition and production with regard to the effect of orthography on L2 word learning. Experiment 1 examined the effect of orthographic information on the learning of a vocalic ([ɔ-u]) free variation, but no effect of orthography on participants' accuracy was found. Both Orth+ and Orth– groups showed low *d'* on the test items, and both L1 groups could not detect the vowel alternations very well, as indicated by the low *d'* scores, which suggest that both L1 groups did not truly learn that [ɔ-u] were in free variation or that [e-a] were contrastive. Experiment 2 was a picture-naming task, which tested participants' spontaneous production of the words they learned in Experiment 1. Exposure to orthography helped participants produce more correct responses as well as more new forms of the free variants and fewer wrong vowels in the minimal pairs. However, no correlation between the performance in Experiment 1 and 2 was found. Considering that in Experiment 1, participants could not detect the vowel alternations very well, Experiment 3 replicated Experiment 1 with consonantal alternations to test whether participants' detection of the sound alternations would improve. Another goal of Experiment 3 was to examine whether the effect of orthography would be stronger given the more consistent grapheme-to-phoneme correspondence of English consonants than vowels. Overall, the effect of orthography was stronger than that in Experiment 1. Yet, the results of Experiment 3 showed an unexpected asymmetry between [p-b] and [t-d] free variation, where exposure to orthography helped learners detect the [p-b] free variation but not [t-d]. Experiment 4, like Experiment 2, was a picture naming task which tested participants' production of the consonant alternations. A benefit of orthography in Taiwanese participants' production of the new forms of the [t-d] free variants was observed, but for the overall production score, exposure to orthography seemed to hinder the production of the free-variant items. In this case, a correlation between Experiment 3 and 4 was found.

The results from all four experiments are summarized in Table 7.1 below. As can be seen from Table 7.1, the effect of orthography on the learning of free variation is inconsistent across the four experiments reported in this study. One general pattern that can be observed though, is the benefit of orthographic exposure on the production of target alternations, especially for Taiwanese participants. On the other hand, in recognition tasks (picture-word matching), the effect of orthography is strongest in American Orth+ on the [p-b] free variation. Other than this effect, the detection on the sound alternations seems to be low in the [t-d] and [ɔ-u] free variation.

Table 7.1. Summary of results from Experiment 1 – 4

Type of Alternations	Experiment	L1 group	Benefit of orthography	Learning of free variation (characterized by *d′* or production of alternating sounds)[10]
[ɔ-u] free variation	1 (picture-word matching)	Taiwanese	Null	low *d′*
		American	Null	low *d′*
	2 (picture naming)	Taiwanese	Yes	Orth+ produced more new forms of free variants than Orth–
		American	Yes	
[p-b] free variation	3 (picture-word matching)	Taiwanese	small effect	low *d′* (although Orth+ > Orth–)
		American	Yes	American Orth+: *d′* = 1.82
	4 (picture naming)	Taiwanese	Yes	Orth+ produced fewer wrong minimal pairs than Orth–
		American	small effect	
[t-d] free variation	3 (picture-word matching)	Taiwanese	Null	low *d′*
		American	Null	low *d′*
	4 (picture naming)	Taiwanese	Yes	Orth+ produced more new forms of free variants than Orth–
		American	Null	No significant difference between Orth+ and Orth–

In the following sections, I will first discuss the learning of free variation as a whole reported in this study and the possible accounts for the limited learning outcomes (section 7.2). Then I will move onto the discussion on the effect of orthography on the learning of free variation and general word learning (section 7.3.1 and 7.3.2). And lastly, I will discuss how the different roles between consonants and vowels in lexical access may account for the difference between the findings from vowel and consonantal alternations (section 7.4).

[10] Note that the production of alternating sounds, which focuses on the target alternations, is different from the production accuracy, which focuses on the prduction of the whole words. Thus, the learning of free variation is reflected by the production of target alternations rather than the production accuracy.

7.2 Learning of Free Variation

The results show that when [ɔ-u] or [t-d] were in free variation, regardless of orthographic conditions, neither L1 group could detect the alternations very well as shown by very low *d’* values. Thus, one might wonder whether participants have really learned the words. As shown by the high *d’* values on the baseline items, they did learn the words and did not just guess on the matched or mismatched items. In other words, both L1 groups showed comparable word learning ability. Therefore, the poor detection of the sound alternation should be attributed to factors other than word learning ability. Several possible accounts for the poor learning of free variation as well as possible future directions are discussed below.

(1) Naturalness of the language/task

In the current design, the items were all nonwords in both English and Mandarin, which were the native languages of the participants. Although all the sounds used in the creation of the experiments are either close to or existent in both English and Mandarin, it is possible that the novelty of the pseudowords might distract participants’ attention from the alternating target sounds. In a study on the learning of morphosyntactic variation, Wonnacott (2011) found that using a *semi-artificial language* where the nouns were real English words and function words were novel words, the degree of regularization over variation (i.e., the tendency to use one form over another) found in child learners was smaller than in other previous studies (e.g., Hudson Kam & Newport, 2005). However, Wonnacott’s study was concerned with morphosyntactic variation and child learners, so it is questionable whether her findings will be generalizable to phonological free variation and adult learners. Considering the generally poor learning of free variation found in the current study, one possibility for future research could be to mix some real English words as distractors with nonwords as test items to make the language less foreign in order to examine whether the learning of free variation would improve when item familiarity is higher.

Secondly, in Experiment 3 and 4, some may argue that the phonology of this artificial language is debatable in that one pair of stops (e.g., [p-b]) is in free variation while another, parallel pair is a minimal pair (e.g., /t-d/). Such a case might be less common typologically, so in order to make the artificial language more typologically plausible, it might be possible to make the voicing contrast occur as free variation such that [t-d] or [p-b] are free variants, while place contrast occurs as minimal pairs (e.g., /c-k/). Then all participants would learn /c-k/ as phonemic, while one group of learners would learn [t-d] as free variants, and another group would learn [p-b] as free variants. If learners were still unable to learn the consonantal free variation very well (as was the case in Experiment 3), the ambiguous status of the voicing feature (i.e., used both in free variation and minimal pairs in Experiment 3) as a possible account for the poor learning of free variation in Experiment 3 (except for American Orth+, see section 5.7.1 for more details) could be ruled out since the voicing feature would never be contrastive in this new design. Such an experiment might provide interesting new directions to probe the possibility of a differential status of certain contrasts for learnability (Peperkamp, Pettinato, & Dupoux, 2003).

(2) Amount of exposure to the free variation

In the current design, each participant was exposed to the rule of free variation (i.e., heard two phonetic variants paired with the same picture) 24 times (6 pairs of free variants times 4 repetitions). It is possible that this amount of exposure is insufficient for learners to fully learn the free variation, so a possibility for future research would be to reduce the number of test items and increase the number of repetitions to increase exposure and meanwhile reduce the task demand. Note that the current design and protocol require about at least an hour to complete, assuming that a given participant needs only one learning cycle. Thus, increasing the number of items as well as the number of repetitions would require even more time and task demand, which may worsen the learning outcome. Balancing constraints of exposure length and participants' fatigue will be challenging but necessary in order to maximize learning outcomes. Another possibility would be to break down the major learning phase into several smaller sessions to reduce the cognitive load, which may potentially improve the learning of free variation.

(3) Motivation

One possible account for learners' poor learning of the free variation (except for American Orth+ on [p-b] free variation) could be their motivation in the task. In the current design, learners were told at the beginning of the familiarization phase that they were not forced to remember every word but to get familiar with the pronunciation of this language. In other words, the familiarization phase was designed to "expose" the learners to the pronunciation and the phonology of this language. It is possible that they might not have learned much of the free variation and minimal pairs if they were not motivated enough to remember the associations between the words and pictures in the familiarization phase. One possible method to increase learner's motivation to learn the words would be to modify the paradigm – it would be possible to convert the familiarization phase as the first part of the learning phase. In other words, participants would be told at the beginning of the first part of the learning phase that there would be a short test following the learning phase, which could increase their motivation to learn and memorize the words.

Another possible way to stimulate participants' motivation for the learning task would be to use incentives, such as monetary reward. For example, in a word learning study conducted by Bürki et al. (2012), participants who scored in top range were awarded 25€ more than those who performed in the bottom range. Using monetary reward for better performance may increase participants' motivation in such a difficult task.

Furthermore, a questionnaire that asks participants to self-evaluate their motivation in L2 word learning (e.g., "Self-Regulating Capacity in Vocabulary Learning" scale developed by Tseng, Dörnyei and Schmitt, 2006) could be helpful in identifying the possible factors accounting for each participants' learning performance.

One demographic factor that might influence learners' motivation is their age and education level. In the present study, most of the American participants were undergraduate students while most of the Taiwanese participants were graduate students. In addition, most of the Taiwanese participants were recruited among personal acquaintances of mine, so that their motivation for the learning task might have been higher than those who were not my acquaintances. Although participants' education level is not necessarily correlated with their performance (e.g., Americans learned the [p-b] free variation better than Taiwanese), it would also be important to control for this factor to rule out any possible confounding effect.

(4) Feedback

It was expected that learners would be able to generalize the alternation patterns they observed in the familiarization and learning phase to the testing phase. The results from Experiment 1 as well as the [t-d] free variation in Experiment 2 suggest that the learning of free variation in both L1 groups was limited. Several other reasons could account for this: 1.) learners did not observe the alternation patterns in the familiarization and learning phase at the first place; or 2.) they did not carry over the learning of free variation to the testing phase due to high task demand. As described in the method sections of Experiment 1 (see section 3.1.1), no feedback was given to participants during the familiarization and the learning phase (even though they were warned that some words may have more than one pronunciation). Thus, in addition to several possible methods to reinforce the learning of free variation discussed above, another possible way would be to add a short test (i.e., picture-word matching) after the familiarization phase and also provide feedback for each trial to build a learning curve for the learning of free variation.

However, the original design of this study was to create a somewhat challenging scenario to examine whether the effect of orthography on word learning would come into play. If feedback had been provided during the familiarization and learning phase, it is possible that learners' detection of the target sound alternations would have significantly improved (i.e., higher *d'* on the test items in the testing phase). In that case, a potential effect of orthography might not be observed, considering that the feedback would have provided an additional aid for learners to learn the free variation, in which case the orthographic information would carry less weight than the feedback. It is clear that more experiments that vary these factors systematically would be needed in order to satisfactorily answer these questions.

7.3 The Effect of Orthography on Word Learning and L2 Phonological Processing

7.3.1 Orthographic effect on general word learning

With respect to the effect of orthography on general word learning – as measured through performance on the baseline items – no clear effect of orthography was found on accuracy, RT and *d'*. This finding suggests that learners were in general quite accurate at word learning overall, regardless of orthography, which is inconsistent with previous studies where orthography was shown to be either facilitative or inhibitory depending on L1-L2 orthography mapping (e.g., Erdener & Burnham, 2005). One possible account for the lack of orthographic effect on general word learning in the current study is that the baseline items did not involve any alternations and thus were easier for learners to learn regardless of their exposure to the orthographic forms. Another possible account is learners' familiarity with the alphabetic writing system. One might assume that Taiwanese participants (Mandarin native speakers), who come from a background of logographic writing system, should show a stronger reliance on the orthographic forms during learning as discussed in the studies of visual word recognition (see Chapter 2. section 2.2.1.3). This could be true if Taiwanese participants had low familiarity with the alphabetic writing system. Note that in the current study, all of the Mandarin speakers (from Taiwan) had intermediate to high level of English proficiency, suggesting a high familiarity with the alphabetic writing system. However, based on the current data, it is not possible to conclude with certainty that the Taiwanese participants in the present study recognized printed words in a way similar to alphabetic readers. In other words, despite their assumed high familiarity with the alphabetic writing system, it is unknown whether their processing of the experimental stimuli was comparable to English native speakers'. An additional experiment tapping into the phonological and orthographic activation during visual word recognition would need to be added to the current protocol to examine whether these participants rely more on orthographic or phonological information during word recognition.

In addition, although Taiwanese participants' familiarity with alphabetic writing system was loosely controlled by their exposure to English-speaking environment (characterized by their length of residence in English-speaking countries), it is possible that Taiwanese participants who have learned other alphabetic languages would be more familiar with alphabetic writing system and thus be influenced by the orthography more than those who have not learned other alphabetic languages. In addition, the stimuli in the current study were matched on the English phonological and orthographic neighborhood, but it is also possible that some words could have had close phonological neighbors in Mandarin, in which case Taiwanese might score higher and/or faster on the items that have close phonological neighbors in Mandarin. So for future research, it would also be necessary to match the stimuli on Mandarin phonological neighborhood. Orthographic neighborhood, however, can only be matched based on English since Mandarin uses a different writing system.

Another possible account for the lack of orthographic effect on general word learning is participants' familiarity with the baseline items. Note that the baseline items in the picture-word matching task were the same as the ones participants learned in the learning phase (as opposed to the test items that appeared as the new forms in the testing phase), and these baseline items also appeared in the criterion test in which they needed to score at least 90% correct. Thus, it is likely that the lack of orthographic effect on general word learning in the current study could be due to the fact that participants had been tested on the same baseline items in the criterion test.

7.3.2 Orthographic effect on the learning of free variation

In Experiments 1 and 3 (where [t-d] were in free variation), the effect of orthography was not strong, as shown by the low *d'* in both L1 groups. The current results hence suggest only partial support for the role of orthography in lexical access – American learners who were exposed to orthography detected the [p-b] free variation significantly better than those who did not see the written forms. Also, in Experiments 2 and 4, although the effect is in the opposite direction (facilitative in Experiment 2 and inhibitory in Experiment 4), orthographic information does seem to influence lexical retrieval and decoding. These findings add further support to the effect of orthography on lexical access – learners' recognition and production of novel words appear to be influenced by the prior presentation of the orthographic forms. However, the present study did not tap into the unconscious processing of orthographic information (e.g., Chéreau et al., 2007), so it is unclear from the current results whether lexical access is mediated by orthographic information or whether orthographic information is stored and accessed at the higher level along with semantic information.

Nevertheless, the current study does contribute to the understanding of the interaction between orthographic and auditory inputs in lexical access. Specifically, the current results add some indirect support to the Bimodal Interactive Activation Model (BIAM) (see Figure 2.3 in section 2.2.2.2) in which the orthographic codes could be activated upon hearing the auditory input even without seeing the written forms. One piece of evidence supporting the BIAM model comes from Experiment 1. As discussed in Chapter 3 (section 3.7.1), one possible account for the poor learning outcome of the vocalic free variation regardless of the exposure to orthography is that the opaque orthography of vowels in English was activated when participants heard the words (regardless of whether they also saw the spellings), and orthographic activation of the vowels might lead to ambiguous recognition of minimal-pair items, which can be characterized by the low *d'* on the vowel alternations, since in English one vowel could be spelled in more than one way. In Experiment 3, when examining consonants, exposure to orthographic forms helped Americans detect the [p-b] free variation, which also supports the BIAM. Nonetheless, it remains unclear why the benefit of orthography was not generalized to other consonantal contrasts: it was not present in the learning of the [t-d] free variation. Thus, the asymmetric effect of orthography between the [p-b] and [t-d] free variation cannot be attributed to the sole influence of orthography on the learning of the free variation.

Despite the null effect of orthography in Experiment 1 and the mixed effect of orthography in Experiment 3, the effect of orthography, although not always facilitative, was also found in production, and a difference between the effect of orthography on production was also found between the naming of pictures containing vocalic and consonantal free variation. In the present study the grapheme-to-phoneme correspondence of the artificial L2 is consistent except for the free variants which were spelled the same (i.e., inconsistent "grapheme-to-sound"[11] correspondence for the free variants). As discussed in Chapter 6, the inconsistent grapheme-to-sound correspondence of the free variants could be the cause for the inhibitory effect of orthography on the production accuracy of the free-variant items. And the difference in the effect of orthography between the naming of the vowel (Experiment 2) and consonant items (Experiment 4) could be due to the orthographic depth of English vowels and consonants. Both L1 groups, Mandarin and English use deep orthography where the grapheme-to-phoneme correspondence is inconsistent, and both L1 groups were familiar with the English orthography. Note that even within the same orthography some sounds could have more transparent

[11] Grapheme-to-sound is used in this context because the free variants are not contrastive and considered to be the realizations of the same phoneme.

grapheme-to-phoneme correspondence than others (e.g., English vowels being more opaque than consonants). As found by Erdener and Burnham (2005), when both L1 and L2 orthography are opaque, small facilitative effects of orthography have been shown to arise. On the other hand, when L1 orthography is transparent and L2 orthography is opaque (shallow orthography), an inhibition effect in L2 word production was found. In the present study, although the orthographic depth for the free variants can be considered as “deep” or opaque, the inconsistent grapheme-to-sound correspondence of the consonants in the current study mismatches with the orthographic depth of English consonants. It is likely that this mismatch between L1 and L2 orthographic depths causes the inhibition on the naming of the free-variant items. Vowels, on the other hand, have inconsistent grapheme-to-phoneme correspondence in English, which coincides with the grapheme-to-sound correspondence of the free variants in the present study. Thus, this might be the reason why exposure to orthography had a facilitative effect on the naming of the vocalic free-variant items.

As found by Erdener and Burnham (2005), the facilitation effect can be even stronger if both L1 and L2 orthography are transparent because the learners are used to the consistent grapheme-to-phoneme correspondence. For future research, in order to further investigate how L1-L2 orthography mapping would influence the effect of orthography on the learning of free variation, L1 groups who use a transparent orthography as well as an opaque L2 orthography will be needed. As shown by Erdener and Burnham (2005), in a production task, Turkish participants from a transparent orthography background made more errors in Irish stimuli (an opaque orthography) than Australian (English) participants when both groups saw the written forms of the words, suggesting an inhibitory effect of orthography when L1 is transparent (Turkish) and L2 is opaque. One possible reason for this effect might be that the orthographic transparency they were expecting from their L1 is transferred to the L2, where it was not appropriate. Participants seemed to expect to draw consistent grapheme-to-phoneme correspondence from the L2 input (even though it is opaque), and it seems that these new correspondences were very difficult to ignore, leading to an inhibitory effect, especially in the production of novel words containing consonant alternations (Experiment 4).

In addition to the orthographic depth, it is also possible to manipulate participants' familiarity with the L2 writing system. For example, in the present study both L1 groups were familiar with alphabet, although Taiwanese participants were L2 speakers of English. A possibility for future research will be to use a script that is familiar to Taiwanese but unfamiliar to English speakers. For instance, in Taiwan, children from kindergarten to 12th grade learn the "Zhu Yin Fu Hao" (or "Bo Po Mo Fo"), a phonetic alphabet for Mandarin developed in the early 20th century (before Pinyin). It might be interesting to use this alphabet to create an L2 orthography to investigate how differences in familiarity with L2 scripts influence the use of orthography in word learning.

7.4 Lexical Access: Vowels vs. Consonants

It was hypothesized that the poor learning of the free variation in Experiment 1 could be attributed to the differential role of consonants and vowels in lexical access in which consonants are tied with lexical access more than vowels (e.g., Bonatti et al., 2005; Cutler et al., 2000), and this effect was expected for both groups equally. However, for the Taiwanese participants, the comparison between the *d'* of Experiments 1 and 3 (see Table 5.10, Chapter 5) showed that these participants did not detect the consonant alternations significantly better than vowel alternations. In fact they detected the vowel alternations slightly better than consonants. On the contrary, Americans detected the consonant alternations better than vowel alternations, although the comparison reached statistical significance only for American Orth+.

Although the comparison of *d'* between Experiments 1 and 3 did not show clear evidence for the stronger role of consonants in lexical access, the bias toward "*no*" response in Experiment 3 (as opposed to the "*yes*" bias in Experiment 1) and the correlation between Experiment 3 (picture-word matching) and 4 (picture naming) might be attributed to the role of consonants in lexical access. As discussed in Chapter 5 (Experiment 3), both L1 groups tended to answer "*no*" to most items (except for Taiwanese Orth– who tended to answer more "*yes*" on [t-d] free variation). It is possible that participants tended to accept only the original forms they learned – they preferred the original forms because of consonants' lesser tendency to vary and their stronger tie with lexical access.

Another possible piece of evidence for the difference between consonants and vowels in lexical access comes from the correlation between Experiments 3 and 4. As discussed in Chapter 6, it is likely that excelling in Experiment 3, which was a picture-word matching for the items of consonant alternations, helped the subsequent picture naming task in Experiment 4. Such a correlation was not found between Experiments 1 and 2. One possible account for the lack of correlation is the weaker role of vowels in lexical access that makes the encoding of the words containing vowel alternations harder, so completing the picture-word matching for the items of vowel alternations did not help picture naming. Yet this speculative account would need to be further confirmed with future studies.

7.5 Conclusion and Perspectives

The present work further contributes to the understanding of the effect of orthography on L2 word learning and phonological acquisition, specifically examining the question of whether orthographic information can help L2 learners link two free variants to one single lexical entry. The results show that exposure to L2 orthography had mixed effects on the learning of free variation. Overall, the results did not fully confirm the benefit of orthography in the learning of free variation, which echoes previous findings about mixed results of the effect of orthography on L2 phonology. Orthographic information did not help the learning of vocalic free variation but helped learners learn the [p-b] free variation. Also, the effect of orthography was stronger in production (picture naming) than in recognition (picture-word matching). In general, both Taiwanese and American participants did not learn the free variation very well (except for the Orth+ group who learned the [p-b] free variation). As discussed earlier, several possible factors that might account for the poor learning of free variation include: the design of the tasks, the naturalness of the target language, learners' motivation, feedback, and so on. Thus, given the possible limitations of the study, a firm conclusion about the learnability of the free variation cannot be made.

In addition, the detectability of the alternations was low in both vowel and consonant alternations, which seemingly disconfirms the differential role of vowels and consonants in lexical access. But the correlation between the performance on picture-word matching (Experiment 3) and picture naming (Experiment 4) seemed to find indirect support for the stronger effect of consonants in lexical access, in that learners might benefit from the previous task on the performance of the subsequent task (picture naming), which is yet to be further investigated with additional experiments.

The present work, to my knowledge, is the first to investigate whether orthographic forms help L2 learners associate pronunciation variants to one single lexical entry. Nevertheless, there still remain several issues that could be further developed into future research, such as the individual differences in the learning of free variation and the potential influence of participants' language background on the effect of orthography. In the following section, I will discuss these two remaining issues as well as some other possible alternative methods to investigate the effect of orthography on word learning and phonological processing.

(1) Individual differences and correlation of cognitive abilities

As observed in the data of Experiment 1, large individual differences were observed on the detection of free variants and minimal pairs (see Figures 3.7 and 3.8). Although it was shown that Taiwanese participants' length of residence positively correlated with their detection of the vowel alternations when they were exposed to the written forms of the words during the familiarization and learning phase, several other factors, such as reading ability, phonological awareness, motivation, cognitive abilities, could underlie the observed differences among individual learner's performance. The sources of individual differences in the present study cannot be confirmed at the current stage. In addition, the design chosen in Experiments 1 and 3 resembled a paired-associate learning task. In Gathercole and Masoura (2003), they found that if second-language learners had higher degree of familiarity with the lexicon of the target L2, phonological short-term memory capacity (characterized by their performance in a nonword repetition task) was not be a significant predictor of L2 learners' performance on paired-associate task which requires participants to learn new words. In the present study, however, all participants had no prior knowledge of the artificial language, so we predict that their performance of word learning might be correlated with working memory measures as well as phonological short-term memory (e.g., Papagno and Vallar, 1995). Thus, future research would benefit from obtaining at least some measures of cognitive abilities (e.g., visual and auditory working memory, attention control, and executive function) in order to examine how those variables correlate with individual learner's performance.

(2) Learners' language background

In the current study, not all American participants were monolingual speakers of English – many of them had learned other foreign languages, which could potentially confound their learning of the free variation depending on their language background, and all Taiwanese participants spoke at least Mandarin and English. Also, learners' familiarity with the alphabetic writing system could also influence the effect of orthography on their word learning.

Given the limited availability of Taiwanese participants at Indiana University who met the length of residence criterion set in the current study, recruiting Taiwanese who are only proficient in Mandarin and English would not be feasible given the time constraint and limited funding. Furthermore, for an even stricter control for Taiwanese participants' language background, in addition to limiting their L2 to English only, their exposure and oral proficiency in Taiwanese (i.e., a variety of Southern Min Chinese) also need to be considered. Most Taiwanese participants were also exposed to or spoke some Taiwanese, which could also be a potential confound in terms of the phonological neighborhood (i.e., some words might sound like a word in Taiwanese, which was not checked in the current study).

(3) Contextual cues

The current paradigm requires learners to learn the new words in isolation. One question is whether the lack of contextual cues made it difficult for the learners to encode the words in their memory, since when learning words of a second language in the real world, learners would learn new words both in isolation and in sentential (or even semantic/situational) contexts. For example, in a perceptual training study done by Greenspan, Nusbaum, and Pisoni (1988), learners benefited from listening to sentences in a series of tasks, compared to learners who were only trained on words in isolation. Nagy, Anderson, and Herman (1987) also found that reading in contexts (passages) helped students learn the meaning of the words. Consequently, one possible direction for future research would be to use the current stimuli to construct several simple sentences accompanied by English translations so that learners have the syntactic and semantic contexts to facilitate their learning.

(4) Context-free vs. context-dependent allophonic variation

One goal of the present study was to investigate whether learners could link two free variants to one lexical entry. Context-free allophonic variation was exploited in order to simplify the design. However, another interesting area for investigation is the context-dependent or position-specific alternation/variation, and how exposure to orthography helps (or inhibits) the learning of context-dependent alternations. For example, the English /p/ is realized as an unaspirated voiceless [p] after /s/ in onset clusters (Kenstowicz, 1994: 258). Although the phonetic realization of /p/ is context-dependent, it is always written as <p> regardless of the surface forms. In this case, one question would be whether exposure to orthography would help L2 learners learn that [p] and [p^h] are allophones of /p/ in

English. Another related issue is whether the presentation of the written forms in fact causes non-native like production (see Bassetti, 2008 for a review). Take /p/ for example. It is likely that L2 learners of English might mistakenly produce [p^h] in "*sport*" due to their exposure to the written forms. All in all, studying how exposure to orthography influences the learning of context-dependent variation may allow us to better understand what resources learners use to learn phonological patterns from the distribution, and also the potential effect of orthography on the production of allophones.

(5) Neurological mechanisms in the processing of free variation and orthographic information

In addition to the behavioral experiments reported in the current thesis, conducting event-related potential (ERP) experiments would allow us to examine the online activation of orthographic information in the learning of free variation. Although in the present study the effect of orthography did not always arise in word learning and the learning of free variation, it is still possible that learners did pay attention to the orthographic form and activate it while doing the picture-word matching task. In the current study, the original forms of the words could be considered as "neutral", while the new forms could be considered as "deviant" of the words they learned in the learning phase. Mismatch negativity (MMN) (Näätänen, Gaillard, & Mäntysalo, 1978) should be observed when listeners hear the new forms. However, if learners have learned the free variation, MMN observed in the new forms of the "free variants" will be smaller than that of the minimal pair items (i.e., new forms of free variant are less strange than a mismatched minimal-pair item). Moreover, Orth− will show larger MMN amplitude for both the new form of the free variant and minimal-pair (i.e., poor learning of the free variation) than Orth+. In other words, Orth+ will show smaller MMN than Orth− because of their better learning of the free variation. In sum, ERP studies will be helpful to better understand the online activation of the orthographic information, in that the timing at which orthography comes into play during lexical access (i.e., prelexical or postlexical) still remains debatable.

Listed above are only some of the remaining issues yet to be resolved and further investigated. As can be seen from both the previous and current research, the effect of orthography on word learning and phonological acquisition may turn out to be quite different in different tasks or different types of contrasts. The present work further contributes to the understanding of the effect of orthography on speech processing and L2 word learning, yet more research needs to be done in order to gain a clearer picture of the role of orthographic representation in phonological encoding as well as lexical representation.

References

Aaronson, D., & Ferres, S. (1986). Sentence processing in Chinese-American bilinguals. *Journal of Memory and Language, 25*(2), 136-162. doi: http://dx.doi.org/10.1016/0749-596X(86)90026-4

Aitchison, J. (2003). *A glossary of language and mind*: Oxford University Press.

Alario, F. X., & Ferrand, L. (1999). A set of 400 pictures standardized for French: Norms for name agreement, image agreement, familiarity, visual complexity, image variability, and age of acquisition. *Behavior Research Methods, Instruments, & Computers, 31*(3), 531-552. doi: 10.3758/BF03200732

Alario, F. X., Perre, L., Castel, C., & Ziegler, J. C. (2007). The role of orthography in speech production revisited. *Cognition, 102*(3), 464-475. doi: 10.1016/j.cognition.2006.02.002

Allport, D. A., & Funnell, E. (1981). Components of the Mental Lexicon. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences, 295*(1077), 397-410. doi: 10.1098/rstb.1981.0148

Bassetti, B. (2008). Orthographic input and second language phonology. In T. Piske & M. Young-Scholten (Eds.), *Input Matters in SLA* (pp. 191-206). Clevedon, UK: Multilingual Matters.

Bassetti, B., & Atkinson, N. (2015). Effects of orthographic forms on pronunciation in experienced instructed second language learners. *Applied Psycholinguistics, 36*(01), 67-91. doi: doi:10.1017/S0142716414000435

Best, C. T. (1994). The emergence of native-language phonological influences in infants: A perceptual assimilation model. In J. Goodman & H. Nusbaum (Eds.), *The development of* *speech perception: The transition from speech sounds to spoken words* (pp. 167-224). Cambridge, MA: MIT Press.

Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In O.-S. Bohn & M. J. Munro (Eds.), *Language experience in second language speech learning: In honor of James Emil Flege* (pp. 13-34). Philadelphia, PA: John Benjamins Publishing.

Boersma, P., & Weenik, D. (2012). Praat: doing phonetics by computer (Version 5.3.11). Retrieved from http://www.praat.org

Bonatti, L. L., Peña, M., Nespor, M., & Mehler, J. (2005). Linguistic Constraints on Statistical Computations: The Role of Consonants and Vowels in Continuous Speech Processing. *Psychological Science, 16*(6), 451-459. doi: 10.1111/j.0956-7976.2005.01556.x

Bradlow, A. R., Akahane-Yamada, R., Pisoni, D. B., & Tohkura, Y. (1999). Training Japanese listeners to identify English/r/and/l: Long-term retention of learning in perception and production. *Perception & Psychophysics, 61*(5), 977-985.

Bradlow, A. R., Pisoni, D. B., Akahane-Yamada, R., & Tohkura, Y. (1997). Training Japanese listeners to identify English/r/and/l: IV. Some effects of perceptual learning on speech production. *The Journal of the Acoustical Society of America, 101*(4), 2299-2310.

Brady, S., Shankweiler, D., & Mann, V. (1983). Speech perception and memory coding in relation to reading ability. *Journal of Experimental Child Psychology, 35*(2), 345-367.

Buckley, E. (2001). *Polish o-Raising and phonological explanation.* Paper presented at the annual meeting of the Linguistic Society of America, Washington, DC.

Bürki, A., Alario, F. X., & Frauenfelder, U. H. (2011). Lexical representation of phonological variants: Evidence from pseudohomophone effects in different regiolects. *Journal of Memory and Language, 64*(4), 424-442. doi: http://dx.doi.org/10.1016/j.jml.2011.01.002

Bürki, A., Spinelli, E., & Gaskell, G. M. (2012). A written word is worth a thousand spoken words: The influence of spelling on spoken-word production. *Journal of Memory and Language, 67*(4), 449-467. doi: http://dx.doi.org/10.1016/j.jml.2012.08.001

Burnham, D. (2003). Language specific speech perception and the onset of reading. *Reading and Writing, 16*(6), 573-609. doi: 10.1023/A:1025593911070

Calvert, G. A., Bullmore, E. T., Brammer, M. J., Campbell, R., Williams, S. C. R., McGuire, P. K., . . . David, A. S. (1997). Activation of Auditory Cortex During Silent Lipreading. *Science, 276*(5312), 593-596. doi: 10.1126/science.276.5312.593

Carroll, S., & Shea, C. (2007). Salience, focus and second language acquisition. *Nouveaux Cahiers de Linguistique Française, 28,* 99-106.

Chao, K.-Y., & Chen, L.-M. (2008). A cross-linguistic study of voice onset time in stop consonant productions. *Computational Linguistics and Chinese Language Processing, 13*(2), 215-232.

Chéreau, C., Gaskell, M. G., & Dumay, N. (2007). Reading spoken words: Orthographic effects in auditory priming. *Cognition, 102*(3), 341-360. doi: http://dx.doi.org/10.1016/j.cognition.2006.01.001

Cho, J.-R., & Chen, H.-C. (1999). Orthographic and Phonological Activation in the Semantic Processing of Korean Hanja and Hangul. *Language and Cognitive Processes, 14*(5-6), 481-502. doi: 10.1080/016909699386167

Clopper, C. G., & Pisoni, D. B. (2008). Perception of Dialect Variation *The Handbook of Speech Perception* (pp. 312-337): Blackwell Publishing Ltd.

Coenen, E., Zwitserlood, P., & Bölte, J. (2001). Variation and assimilation in German: Consequences for lexical access and representation. *Language and Cognitive Processes, 16*, 535-564.

Coltheart, M. (1981). The MRC psycholinguistic database. *The Quarterly Journal of Experimental Psychology Section A, 33*(4), 497-505. doi: 10.1080/14640748108400805

Connine, C., M., & Pinnow, E. (2006). Phonological variation in spoken word recognition: Episodes and abstractions *The Linguistic Review* (Vol. 23, pp. 235).

Cutler, A., Sebastián-Gallés, N., Soler-Vilageliu, O., & Ooijen, B. (2000). Constraints of vowels and consonants on lexical selection: Cross-linguistic comparisons. *Memory & Cognition, 28*(5), 746-755. doi: 10.3758/BF03198409

Cutler, A., Weber, A., & Otake, T. (2006). Asymmetric mapping from phonetic to lexical representations in second-language listening. *Journal of Phonetics, 34*(2), 269-284. doi: 10.1016/j.wocn.2005.06.002

Damian, M. F., & Bowers, J. S. (2003). Effects of orthography on speech production in a form-preparation paradigm. *Journal of Memory and Language, 49*(1), 119-132. doi: 10.1016/s0749-596x(03)00008-1

Darcy, I., Dekydtspotter, L., Sprouse, R. A., Glover, J., Kaden, C., McGuire, M., & Scott, J. H. (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*(1), 5-40. doi: 10.1177/0267658311423455

Davidson, L. (2007). *The influence of non-phonetic factors on the form of L2 lexical entries: response to Cutler and Weber.* Paper presented at the ICPhS XVI, Saarbrücken, Germany.

de Jong, K., & Park, H. (2012). Vowel epenthesis and segment identity in Korean learners of English. *Studies in Second Language Acquisition, 34*(01), 127-155. doi: doi:10.1017/S0272263111000520

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

Detey, S., & Nespoulous, J.-L. (2008). Can orthography influence second language syllabic segmentation?: Japanese epenthetic vowels and French consonantal clusters. *Lingua, 118*(1), 66-81. doi: 10.1016/j.lingua.2007.04.003

Diependaele, K., Ziegler, J. C., & Grainger, J. (2010). Fast phonology and the bimodal interactive activation model. *European Journal of Cognitive Psychology, 22*(5), 764-778.

Dijkstra, T., Roelofs, A., & Fieuws, S. (1995). Orthographic effects on phoneme monitoring. *Canadian Journal of Experimental Psychology, 49*(2), 264.

Dijkstra, T., & Van Heuven, W. J. B. (1998). The BIA model and bilingual word recognition. In J. Grainger & A. M. Jacobs (Eds.), *Localist connectionist approaches to human cognition* (pp. 189-225). Mahwah, NJ: Lawrence Erlbaum Associates.

Dijkstra, T., & Van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition, 5*(03), 175-197. doi: 10.1017/S1366728902003012

Dornbusch, T. (2012). *Orthographic influences on L2 auditory word processing.* (Ph.D. Dissertation), Technische Universität Dortmund.

Dupoux, E., Kakehi, K., Hirose, Y., Pallier, C., & Mehler, J. (1999). Epenthetic vowels in Japanese: A perceptual illusion? *Journal of Experimental Psychology: Human perception and performance, 25*(6), 1568-1578.

Dupoux, E., & Mehler, J. (1992). Unifying awareness and on-line studies of speech: A tentative framework. In D. H. J. Alegria, J. Junca de Morais and M. Radeau (Ed.), *Analytic Approaches to Human Cognition* (pp. 59-75). Oxford, England: Elsevier Publishers.

Dupoux, E., Pallier, C., Kakehi, K., & Mehler, J. (2001). New evidence for prelexical phonological processing in word recognition. *Language and Cognitive Processes, 16*(5-6), 491-505. doi: 10.1080/01690960143000191

Dupoux, E., Pallier, C., Sebastian, N., & Mehler, J. (1997). A destressing “deafness” in French? *Journal of Memory and Language, 36*(3), 406-421.

Dupoux, E., Peperkamp, S., & Sebastián-Gallés, N. (2001). A robust method to study stress “deafness”. *The Journal of the Acoustical Society of America, 110*, 1606.

Dupoux, E., Sebastián-Gallés, N., Navarrete, E., & Peperkamp, S. (2008). Persistent stress ‘deafness’: The case of French learners of Spanish. *Cognition, 106*(2), 682-706.

Ehrich, J. F., Zhang, L. J., Mu, J. C., & Ehrich, L. C. (2012). Are alphabetic language-derived models of L2 reading relevant to L1 logographic background readers? *Language Awareness, 22*(1), 39-55. doi: 10.1080/09658416.2011.644796

Elman, J. L. (2004). An alternative view of the mental lexicon. *Trends in Cognitive Sciences, 8*(7), 301-306. doi: http://dx.doi.org/10.1016/j.tics.2004.05.003

Erdener, V. D., & Burnham, D. K. (2005). The Role of Audiovisual Speech and Orthographic Information in Nonnative Speech Production. *Language Learning, 55*(2), 191-228. doi: 10.1111/j.0023-8333.2005.00303.x

Escudero, P. (2015). Orthography plays a limited role when learning the phonological forms of new words: The case of Spanish and English learners of novel Dutch words. *Applied Psycholinguistics, 36*(01), 7-22. doi: doi:10.1017/S014271641400040X

Escudero, P., Hayes-Harb, R., & Mitterer, H. (2008). Novel second-language words and asymmetric lexical access. *Journal of Phonetics, 36*(2), 345-360. doi: http://dx.doi.org/10.1016/j.wocn.2007.11.002

Escudero, P., Simon, E., & Mulak, K. E. (2014). Learning words in a new language: Orthography doesn't always help. *Bilingualism: Language and Cognition, 17*(02), 384-395. doi: doi:10.1017/S1366728913000436

Escudero, P., & Wanrooij, K. (2010). The effect of L1 orthography on non-native vowel perception. *Language and Speech, 53*(3), 343-365. doi: 10.1177/0023830910371447

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233-277). Timonium, MD: York Press.

Fromkin, V. A. (1987). The Lexicon: Evidence from Acquired Dyslexia. *Language, 63*(1), 1-22. doi: 10.2307/415382

Frost, R., & Ziegler, J. C. (2007). Speech and spelling interaction: the interdependence of visual and auditory word recognition. In G. Gaskell (Ed.), *The Oxford handbook of psycholinguistics* (pp. 107-118). New York: Oxford University Press.

Fry, E. (2004). Phonics: A Large Phoneme - Grapheme Frequency Count Revised. *Journal of Literacy Research, 36*(1), 85-98. doi: 10.1207/s15548430jlr3601_5

Gaskell, M. G. (2007). Statistical and connectionist models of speech perception and word recognition. In G. Gaskell (Ed.), *The Oxford handbook of psycholinguistics* (pp. 55-70). New York: Oxford University Press.

Gaskell, M. G., & Marslen-Wilson, W. D. (1997). Integrating Form and Meaning: A Distributed Model of Speech Perception. *Language and Cognitive Processes, 12*(5-6), 613-656. doi: 10.1080/016909697386646

Gathercole, S. E., & Masoura, E. V. (2003). Contrasting contributions of phonological short-term memory and long-term knowledge to vocabulary learning in a foreign language. *Memory, 13*(3-4), 422-429. doi: 10.1080/09658210344000323

Gourevitch, V., & Galanter, E. (1967). A significance test for one parameter isosensitivity functions. *Psychometrika, 32*(1), 25-33. doi: 10.1007/BF02289402

Grainger, J., & Ferrand, L. (1994). Phonology and Orthography in Visual Word Recognition: Effects of Masked Homophone Primes. *Journal of Memory and Language, 33*(2), 218-233. doi: http://dx.doi.org/10.1006/jmla.1994.1011

Grainger, J., & Ferrand, L. (1996). Masked Orthographic and Phonological Priming in Visual Word Recognition and Naming: Cross-Task Comparisons. *Journal of Memory and Language, 35*(5), 623-647. doi: http://dx.doi.org/10.1006/jmla.1996.0033

Grainger, J., & Ziegler, J. (2011). A dual-route approach to orthographic processing. *Frontiers in psychology, 2*. doi: 10.3389/fpsyg.2011.00054

Grainger, J., & Ziegler, J. C. (2008). Cross-code consistency in a functional architecture for word recognition. In E. L. Grigorenko & A. J. Naples (Eds.), *Single-word reading: behavioral and biological perspectives* (pp. 129-157). New York, NY: Lawrence Erlbuam Associates.

Greenspan, S. L., Nusbaum, H. C., & Pisoni, D. B. (1988). Perceptual Learning of Synthetic Speech Produced by Rule. *Journal of experimental psychology. Learning, memory, and cognition, 14*(3), 421-433.

Hanna, P. R., Hanna, J. S., Hodges, R. E., & Rudorf, E. H. (1966). *Phoneme-grapheme correspondences as cues to spelling improvemtn*. Washington, DC: U.S. Department of Health, Education, and Welfare.

Hayes-Harb, R., Nicol, J., & Barker, J. (2010). Learning the Phonological Forms of New Words: Effects of Orthographic and Auditory Input. *Language and Speech, 53*(3), 367-381. doi: 10.1177/0023830910371460

Huang, H. S., & Hanley, J. R. (1995). Phonological awareness and visual skills in learning to read Chinese and English. *Cognition*, *54*(1), 73-98. doi: http://dx.doi.org/10.1016/0010-0277(94)00641-W

Hudson Kam, C. L., & Chang, A. (2009). Investigating the cause of language regularization in adults: Memory constraints or learning effects? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*(3), 815-821. doi: 10.1037/a0015097

Hudson Kam, C. L., & Newport, E. L. (2005). Regularizing Unpredictable Variation: The Roles of Adult and Child Learners in Language Formation and Change. *Language Learning and Development*, *1*(2), 151-195. doi: 10.1080/15475441.2005.9684215

Jakimik, J., Cole, R. A., & Rudnicky, A. I. (1985). Sound and spelling in spoken word recognition. *Journal of Memory and Language, 24*(2), 165-178. doi: 10.1016/0749-596x(85)90022-1

Joanisse, M. F., Manis, F. R., Keating, P. A., & Seidenberg, M. S. (2000). Language Deficits in Dyslexic Children: Speech Perception, Phonology, and Morphology. *Journal of*

Experimental Child Psychology, 77(1), 30-60. doi: http://dx.doi.org/10.1006/jecp.1999.2553

Katz, L., & Frost, R. (1992). Chapter 4 The Reading Process is Different for Different Orthographies: The Orthographic Depth Hypothesis. In F. Ram & K. Leonard (Eds.), *Advances in Psychology* (Vol. Volume 94, pp. 67-84): North-Holland.

Kaushanskaya, M., & Marian, V. (2009). Bilingualism reduces native-language interference during novel-word learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(3), 829-835. doi: 10.1037/a0015275

Keating, P. A. (1991). Coronal places of articulation. In C. Paradis & J.-F. Prunet (Eds.), *The Special Status of Coronals: Internal and External Evidence* (pp. 29-48): Academic Press.

Kemps, R., Ernestus, M., Schreuder, R., & Baayen, H. (2004). Processing reduced word forms: The suffix restoration effect. *Brain and Language, 90*(1–3), 117-127. doi: http://dx.doi.org/10.1016/S0093-934X(03)00425-5

Kenstowicz, M. J. (1994). *Phonology in generative grammar*: Blackwell Cambridge, MA.

Kerzel, D., & Bekkering, H. (2000). Motor activation from visible speech: evidence from stimulus response compatibility. *Journal of experimental psychology. Human perception and performance, 26*(2), 634-647. doi: 10.1037//0096-1523.26.2.634

Klatt, D. H. (1979). Speech perception: A model of acoustic-phonetic analysis and lexical access. *Journal of Phonetics*, *7*(312), 1-26.

Koda, K. (1989). Effects of L1 orthographic representation on L2 phonological coding strategies. *Journal of Psycholinguistic Research, 18*(2), 201-222. doi: 10.1007/bf01067782

Kouider, S., & Dupoux, E. (2001). A functional disconnection between spoken and visual word recognition: evidence from unconscious priming. *Cognition*, *82*(1), B35-B49. doi: http://dx.doi.org/10.1016/S0010-0277(01)00152-4

Kraljic, T., Brennan, S. E., & Samuel, A. G. (2008). Accommodating variation: Dialects, idiolects, and speech processing. *Cognition, 107*(1), 54-81. doi: http://dx.doi.org/10.1016/j.cognition.2007.07.013

Kroll, J. F., & Sunderman, G. (2003). Cognitive processes in second language learners and bilinguals: The development of lexical and conceptual representations. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 104-129). Malden, MA: Blackwell Publishing.

Kuhl, P. K. (1991). Human adults and human infants show a "perceptual magnet effect" for the prototypes of speech categories, monkeys do not. *Perception & Psychophysics, 50*(2), 93-107. doi: 10.3758/BF03212211

Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science, 255*(5044), 606-608. doi: 10.1126/science.1736364

Ladefoged, P. (2000). *A course in phonetics*. Boston, MA: Thomson Wadsworth.

Ladefoged, P. (2004). Index of sounds. Retrieved 8/7, 2014, from http://www.phonetics.ucla.edu/index/sounds.html

Lahiri, A., & Marslen-Wilson, W. (1991). The mental representation of lexical form: A phonological approach to the recognition lexicon. *Cognition, 38*(3), 245-294. doi: 10.1016/0010-0277(91)90008-R

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological review, 74*(6), 431-461. doi: 10.1037/h0020279

Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition, 21*(1), 1-36.

Lisker, L., & Abramson, A. S. (1963). Crosslanguage study of voicing in initial stops. *The Journal of the Acoustical Society of America, 35*, 1889.

Lively, S. E., Logan, J. S., & Pisoni, D. B. (1993). Training Japanese listeners to identify English/r/and/l/. II: The role of phonetic environment and talker variability in learning new perceptual categories. *The Journal of the Acoustical Society of America, 94*(3), 1242-1255.

Logan, J. S., Lively, S. E., & Pisoni, D. B. (1991). Training Japanese listeners to identify English/r/and/l: A first report. *The Journal of the Acoustical Society of America, 89*(2), 874-886.

Luce, P. A., Goldinger, S. D., Auer, E. T., & Vitevitch, M. S. (2000). Phonetic priming, neighborhood activation, and PARSYN. *Perception & Psychophysics, 62*(3), 615-625. doi: 10.3758/BF03212113

Luce, P. A., & McLennan, C. T. (2008). Spoken Word Recognition: The Challenge of Variation *The Handbook of Speech Perception* (pp. 590-609): Blackwell Publishing Ltd.

Luce, P. A., & Pisoni, D. B. (1998). Recognizing Spoken Words: The Neighborhood Activation Model. *Ear and Hearing, 19*(1), 1-36.

Luche, C. D., Poltrock, S., Goslin, J., New, B., Floccia, C., & Nazzi, T. (2014). Differential processing of consonants and vowels in the auditory modality: A cross-linguistic study. *Journal of Memory and Language, 72*, 1-15. doi: 10.1016/j.jml.2013.12.001

Lukatela, G., Frost, S. J., & Turvey, M. T. (1998). Phonological Priming by Masked Nonword Primes in the Lexical Decision Task. *Journal of Memory and Language, 39*(4), 666-683. doi: 10.1006/jmla.1998.2599

Lukatela, G., & Turvey, M. T. (1994). Visual lexical access is initially phonological: I. Evidence from associative priming by words, homophones, and pseudohomophones. *Journal of Experimental Psychology: General, 123*(2), 107-128.

Luo, C. R., Johnson, R. A., & Gallo, D. A. (1998). Automatic activation of phonological information in reading: Evidence from the semantic relatedness decision task. *Memory & Cognition, 26*(4), 833-843.

Manis, F. R., McBride-Chang, C., Seidenberg, M. S., Keating, P., Doi, L. M., Munson, B., & Petersen, A. (1997). Are Speech Perception Deficits Associated with Developmental Dyslexia? *Journal of Experimental Child Psychology, 66*(2), 211-235. doi: http://dx.doi.org/10.1006/jecp.1997.2383

Marslen-Wilson, W., & Warren, P. (1994). Levels of perceptual representation and process in lexical access: words, phonemes, and features. *Psychological review, 101*(4), 653.

Massaro, D. W., & Jesse, A. (2007). Audiovisual speech perception and word. In G. Gaskell (Ed.), *The Oxford handbook of psycholinguistics* (pp. 19-36). New York: Oxford University Press.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology, 18*(1), 1-86. doi: http://dx.doi.org/10.1016/0010-0285(86)90015-0

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological review, 88*(5), 375.

McGurk, H., & Macdonald, J. (1976). Hearing lips and seeing voices. *Nature, 264*(5588), 746-748.

McKeown, M. G., & Curtis, M. E. (Eds.). (1987). *The nature of vocabulary acquisition*: Lawrence Erlbaum Associates, Inc.

McLennan, C. T., Luce, P. A., & Charles-Luce, J. (2003). Representation of lexical form. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*(4), 539 - 553.

McQueen, J. M. (2007). Eight questions about spoken word recognition. In G. Gaskell (Ed.), *The Oxford handbook of psycholinguistics* (pp. 37-54). New York: Oxford University Press.

McQueen, J. M., Cutler, A., & Norris, D. (2006). Phonological Abstraction in the Mental Lexicon. *Cognitive Science, 30*(6), 1113-1126. doi: 10.1207/s15516709cog0000_79

Mody, M., Studdert-Kennedy, M., & Brady, S. (1997). Speech Perception Deficits in Poor Readers: Auditory Processing or Phonological Coding? *Journal of Experimental Child Psychology, 64*(2), 199-231. doi: http://dx.doi.org/10.1006/jecp.1996.2343

Näätänen, R., Gaillard, A. W. K., & Mäntysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta Psychologica, 42*(4), 313-329. doi: http://dx.doi.org/10.1016/0001-6918(78)90006-9

Nagy, W. E., Anderson, R. C., & Herman, P. A. (1987). Learning Word Meanings From Context During Normal Reading. *American Educational Research Journal, 24*(2), 237-270. doi: 10.3102/00028312024002237

New, B., Araújo, V., & Nazzi, T. (2008). Differential Processing of Consonants and Vowels in Lexical Access Through Reading. *Psychological Science, 19*(12), 1223-1227. doi: 10.1111/j.1467-9280.2008.02228.x

Norris, D. (1994). Shortlist: a connectionist model of continuous speech recognition. *Cognition, 52*(3), 189-234. doi: http://dx.doi.org/10.1016/0010-0277(94)90043-4

Nusbaum, H. C., Pisoni, D. B., & Davis, C. K. (1984). Sizing up the Hoosier mental lexicon: Measuring the familiarity of 20,000 words. *Research on Speech Perception Progress Report, 10,* 357-376.

Öhman, S. E. G. (1966). Coarticulation in VCV Utterances: Spectrographic Measurements. *The Journal of the Acoustical Society of America, 39*(1), 151-168. doi: doi:10.1121/1.1909864

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.

Paivio, A. (1986). Mental representations : a dual coding approach. New York: Oxford University Press.

Paivio, A. (2007). Mind and its evolution: a dual coding theoretical approach. Mahwah, N.J.: L. Erlbaum Associates.

Palmeri, T. J., Goldinger, S. D., & Pisoni, D. B. (1993). Episodic encoding of voice attributes and recognition memory for spoken words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*(2), 309-328. doi: 10.1037/0278-7393.19.2.309

Papagno, C., & Vallar, G. (1995). Verbal Short-term Memory and Vocabulary Learning in Polyglots. *The Quarterly Journal of Experimental Psychology Section A, 48*(1), 98-107. doi: 10.1080/14640749508401378

Paradis, C., & Prunet, J.-F. (1991). Introduction: Asymmetry and visibility in consonant articulatations. In C. Paradis & J.-F. Prunet (Eds.), *The Special Status of Coronals: Internal and External Evidence* (pp. 1-28): Academic Press.

Pattamadilok, C., Knierim, I. N., Kawabata Duncan, K. J., & Devlin, J. T. (2010). How does learning to read affect speech perception? *The Journal of Neuroscience, 30*(25), 8435-8444. doi: 10.1523/jneurosci.5791-09.2010

Peperkamp, S., Pettinato, M., & Dupoux, E. (2003). Allophonic variation and the acquisition of phoneme categories. In B. Beachley, A. Brown & F. Conlin (Eds.), *Proceedings of the 27th Annual Boston University Conference on Language Development* (Vol. 2, pp. 650-661). Sommerville, MA: Cascadilla Press.

Perfetti, C. A., & Zhang, S. (1991). Phonological processes in reading Chinese characters. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*(4), 633-643.

Perrachione, T. K., Lee, J., Ha, L. Y. Y., & Wong, P. C. M. (2011). Learning a novel phonological contrast depends on interactions between individual differences and training paradigm design. *The Journal of the Acoustical Society of America, 130*(1), 461-472. doi: doi:http://dx.doi.org/10.1121/1.3593366

Petrova, A., Gaskell, G., & Ferrand, L. (2011). Orthographic consistency and word-frequency effects in auditory word recognition: New evidence from lexical decision and rime detection. *Frontiers in psychology, 2*. doi: 10.3389/fpsyg.2011.00263

Phillips, B. S. (1999). The Mental Lexicon: Evidence from Lexical Diffusion. *Brain and Language, 68*(1–2), 104-109. doi: http://dx.doi.org/10.1006/brln.1999.2104

Pisoni, D. B. (1993). Long-term memory in speech perception: Some new findings on talker variability, speaking rate and perceptual learning. *Speech Communication, 13*(1–2), 109-125. doi: http://dx.doi.org/10.1016/0167-6393(93)90063-Q

Pisoni, D. B., & Lively, S. E. (1995). Variability and invariance in speech perception: A new look at some old problems in perceptual learning. *Speech perception and linguistic experience: issues in cross-language research (ed. W. Strange)*, 429-455.

Polivanov, E. (1931). La perception des sons d'une langue étrangère. *Travaux du cercle linguistique de Prague, 4,* 79-96.

Port, R. F. (2010). Rich memory and distributed phonology. *Language Sciences, 32*(1), 43-55.

Rafat, Y. (2011). *Orthography-induced Transfer in the Production of Novice Adult English-speaking Learners of Spanish.* (Ph.D. Dissertation), University of Toronto.

Rafat, Y. (2013). Orthography-induced transfer in the production of English-speaking learners of Spanish. *The Language Learning Journal*, 1-17. doi: 10.1080/09571736.2013.784346

Rafat, Y. (2015). The interaction of acoustic and orthographic input in the acquisition of Spanish assibilated/fricative rhotics. *Applied Psycholinguistics, 36*(01), 43-66. doi: doi:10.1017/S0142716414000423

Rastle, K., Harrington, J., & Coltheart, M. (2002). 358,534 nonwords: The ARC Nonword Database. *The Quarterly Journal of Experimental Psychology Section A, 55*(4), 1339-1362. doi: 10.1080/02724980244000099

Read, C., Yun-Fei, Z., Hong-Yin, N., & Bao-Qing, D. (1986). The ability to manipulate speech sounds depends on knowing alphabetic writing. *Cognition, 24*(1), 31-44. doi: 10.1016/0010-0277(86)90003-X

Reetz, H., & Jongman, A. (2011). *Phonetics: transcription, production, acoustics, and perception* (Vol. 34): John Wiley & Sons.

Reilly, J., Troiani, V., Grossman, M., & Wingfield, R. (2007). An introduction to hearing loss and screening procedures for behavioral research. *Behavior research methods*, *39*(3), 667-672.

Rogers, H. (2005). *Writing systems: A linguistic approach*. Malden, MA: Blackwell Publishing.

Rumelhart, D. E., & McClelland, J. L. (1982). An interactive activation model of context effects in letter perception: II. The contextual enhancement effect and some tests and extensions of the model. *Psychological review*, *89*(1), 60.

Scharenborg, O., Norris, D., ten Bosch, L., & McQueen, J. M. (2005). How Should a Speech Recognizer Work? *Cognitive Science*, *29*(6), 867-918. doi: 10.1207/s15516709cog0000_37

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime: User's guide*: Psychology Software Incorporated.

Seidenberg, M. S. (1985). The time course of phonological code activation in two writing systems. *Cognition*, *19*(1), 1-30. doi: http://dx.doi.org/10.1016/0010-0277(85)90029-0

Showalter, C. E. (2012). *The Influence of Novel Orthographic Information on Second Language Word Learning: The Case of Native English Speakers Learning Arabic.* (Master of Arts Thesis), The University of Utah.

Showalter, C. E., & Hayes-Harb, R. (2013). Unfamiliar orthographic information and second language word learning: A novel lexicon study. *Second Language Research*, *29*(2), 185-200. doi: 10.1177/0267658313480154

Showalter, C. E., & Hayes-Harb, R. (2015). Native English speakers learning Arabic: The influence of novel orthographic information on second language phonological acquisition. *Applied Psycholinguistics*, *36*(01), 23-42. doi: doi:10.1017/S0142716414000411

Silveira, R. (2007). *Investigating the role of orthography in the acquisition of L2 pronunciation: A case study*. Paper presented at the New Sounds: Fifth International Symposium on the Acquisition of Second Language Speech, Florianópolis, Santa Catarina, Brazil.

Simon, E., Chambless, D., & Kickhöfel Alves, U. (2010). Understanding the role of orthography in the acquisition of a non-native vowel contrast. *Language Sciences, 32*(3), 380-394. doi: 10.1016/j.langsci.2009.07.001

Sommers, M. S. (2003). Speech & Hearing Lab Neighborhood Database. from http://neighborhoodsearch.wustl.edu/Home.asp

Sprenger-Charolles, L., Siegel, L. S., Béchennec, D., & Serniclaes, W. (2003). Development of phonological and orthographic processing in reading aloud, in silent reading, and in spelling: A four-year longitudinal study. *Journal of Experimental Child Psychology, 84*(3), 194-217. doi: http://dx.doi.org/10.1016/S0022-0965(03)00024-9

Steele, J. (2005). *Assessing the role of orthographic versus uniquely auditory input in acquiring new L2 segments*. Paper presented at the 7èmes rencontres internationales du réseau français de phonologie (RFP 2005).

Steffensen, M. S., Goetz, E. T., & Cheng, X. (1999). A Cross-Linguistic Perspective on Imagery and Affect in Reading: Dual Coding in Chinese and English. *Journal of Literacy Research, 31*(3), 293-319. doi: 10.1080/10862969909548050

Stemberger, J. P., & Stoel-Gammon, C. (1991). The underspeficication of coronals: evidence from language acquisition and performance errors. In C. Paradis & J.-F. Prunet (Eds.), *The Special Status of Coronals: Internal and External Evidence* (pp. 181-199): Academic Press.

Strange, W., Jenkins, J. J., & Johnson, T. L. (1983). Dynamic specification of coarticulated vowels. *The Journal of the Acoustical Society of America, 74*(3), 695-705. doi: dx.doi.org/10.1121/1.389855

Sumby, W. H., & Pollack, I. (1954). Visual Contribution to Speech Intelligibility in Noise. *The Journal of the Acoustical Society of America, 26*(2), 212-215. doi: doi:http://dx.doi.org/10.1121/1.1907309

Swingley, D., & Aslin, R. N. (2000). Spoken word recognition and lexical representation in very young children. *Cognition, 76*(2), 147-166.

Taft, M., Castles, A., Davis, C., Lazendic, G., & Nguyen-Hoan, M. (2008). Automatic activation of orthography in spoken word recognition: Pseudohomograph priming. *Journal of Memory and Language, 58*(2), 366-379. doi: 10.1016/j.jml.2007.11.002

Tseng, W.-T., Dörnyei, Z., & Schmitt, N. (2006). A New Approach to Assessing Strategic Learning: The Case of Self-Regulation in Vocabulary Acquisition. *Applied Linguistics, 27*(1), 78-102. doi: 10.1093/applin/ami046

Vitevitch, M. S., & Luce, P. A. (1998). When Words Compete: Levels of Processing in Perception of Spoken Words. *Psychological Science, 9*(4), 325-329. doi: 10.1111/1467-9280.00064

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.

Warren, P., & Marslen-Wilson, W. (1987). Continuous uptake of acoustic cues in spoken word recognition. *Perception & Psychophysics, 41*(3), 262-275. doi: 10.3758/BF03208224

Weber, A., & Scharenborg, O. (2012). Models of spoken-word recognition. *Wiley Interdisciplinary Reviews: Cognitive Science*, *3*(3), 387-401. doi: 10.1002/wcs.1178

Wonnacott, E. (2011). Balancing generalization and lexical conservatism: An artificial language study with child learners. *Journal of Memory and Language*, *65*(1), 1-14. doi: http://dx.doi.org/10.1016/j.jml.2011.03.001

Young-Scholten, M., & Langer, M. (2015). The role of orthographic input in second language German: Evidence from naturalistic adult learners' production. *Applied Psycholinguistics, 36*(01), 93-114. doi: doi:10.1017/S0142716414000447

Zhou, Y. G. (1978). To what degree are the phonetics of present-day Chinese characters still phonetic? (published in Chinese). *Zhongguo Yuwen*, *146*, 172-177.

Ziegler, J., & Ferrand, L. (1998). Orthography shapes the perception of speech: The consistency effect in auditory word recognition. *Psychonomic Bulletin & Review, 5*(4), 683-689. doi: 10.3758/bf03208845

Ziegler, J., Ferrand, L., & Montant, M. (2004). Visual phonology: The effects of orthographic consistency on different auditory word recognition tasks. *Memory & Cognition*, *32*(5), 732-741. doi: 10.3758/bf03195863

Ziegler, J., Muneaux, M., & Grainger, J. (2003). Neighborhood effects in auditory word recognition: Phonological competition and orthographic facilitation. *Journal of Memory and Language*, *48*(4), 779-793. doi: http://dx.doi.org/10.1016/S0749-596X(03)00006-8

Appendix I. Items used in Experiment 1

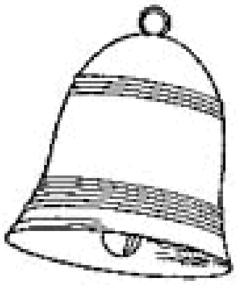
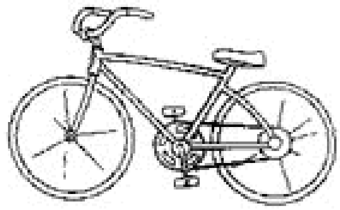
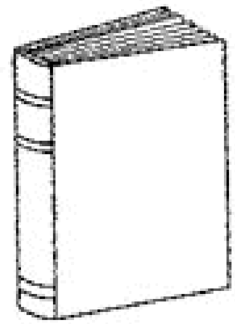
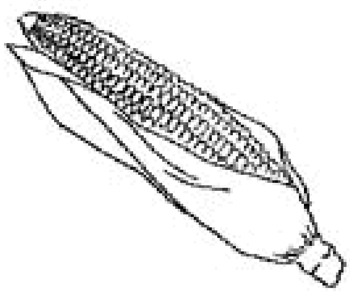
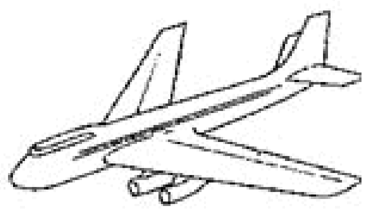
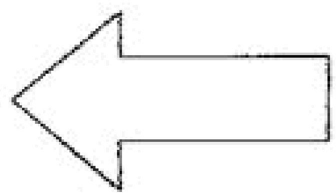
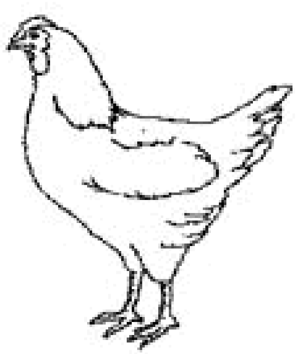
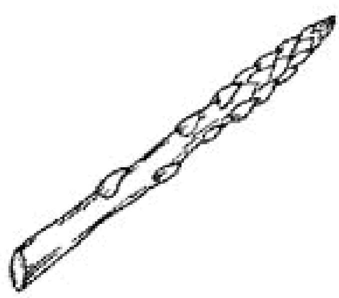
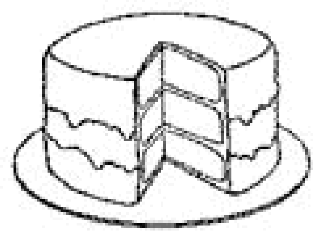
Phase	Minimal Pairs	Free Variants	Baseline
Familiarization	bapip	fopip	fivep
	bepip	fupip	kinis
	falep	mogal	ligik
	felep	mohif	nakel
	masif	mugal	sevek
	mesif	muhif	tetif
	nasip	pogak	
	nesip	pugak	
Learning and Testing	fahal	bofes	divat
	fehal	bufes	fivis
	gakaf	dodet	gefit
	gekaf	dudet	lasis
	nehas	fusat	medip
	nahas	fosat	panek
	sabip	mohil	sesal
	sebip	muhil	tiges
	tatef	solif	
	tetef	sulif	
	tehet	tufap	
	tahet	tofap	

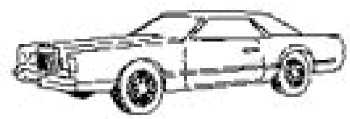
Appendix II. Items used in Experiment 3

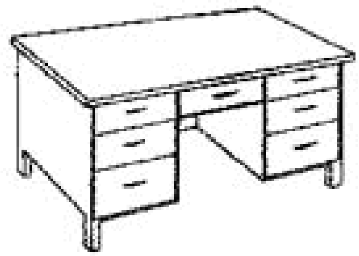
Phase	Minimal Pairs	Free Variants	Baseline
Familiarization	dakuf	bahik	finok
	takuf	pahik	hukel
	deluk	bemaf	kafus
	teluk	pemaf	kosal
	dilof	bomek	lagek
	tilof	pomek	mulil
	duhel	bukol	
	tuhel	pukol	
Learning and Testing	damik	balof	honal
	tamik	palof	kekul
	dehol	bekus	kogaf
	tehol	pekus	lamok
	disef	biluk	liguk
	tisef	piluk	nekis
	dolak	bimal	nelik
	tolak	pimal	sihof
	dokil	bofel	
	tokil	pofel	
	dumak	bukis	
	tumak	pukis	

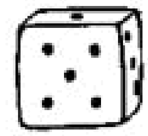
Appendix III. Pictures used in all four experiments

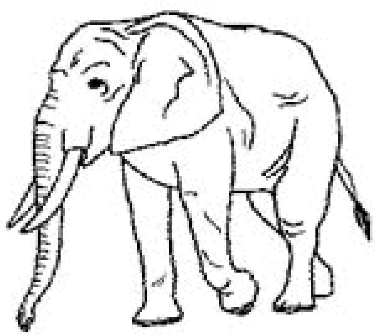
Pictures were selected from Alario and Ferrand. (1999).

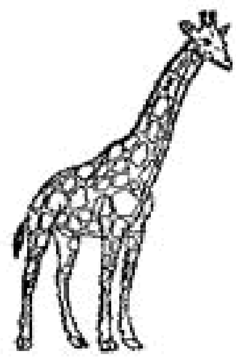










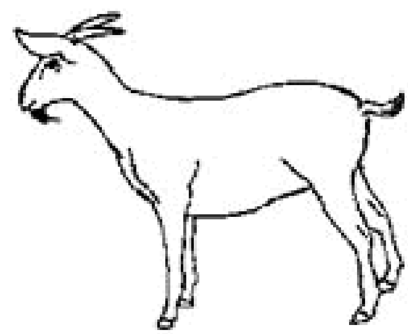


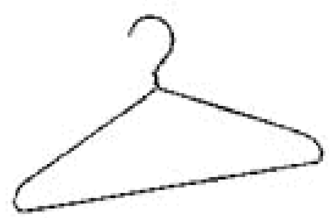


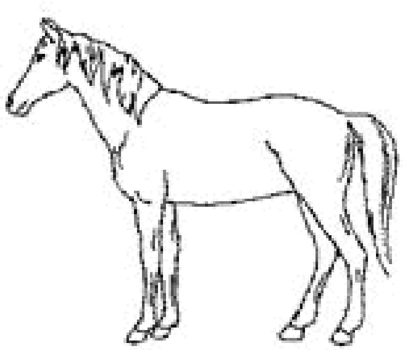


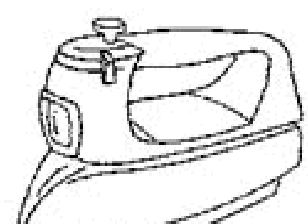


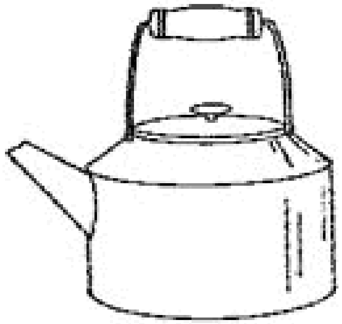
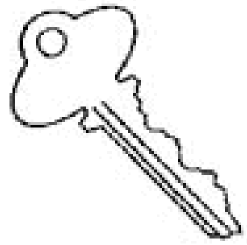
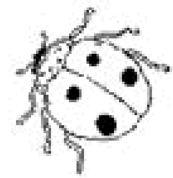
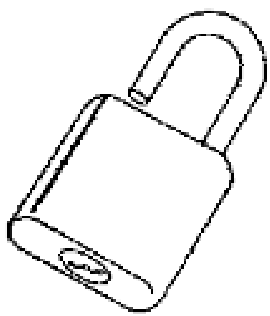
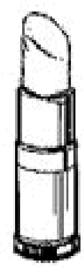
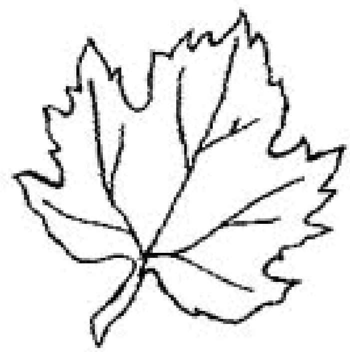


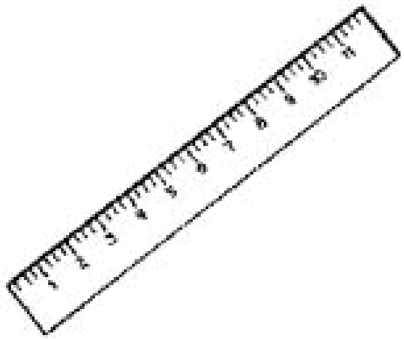
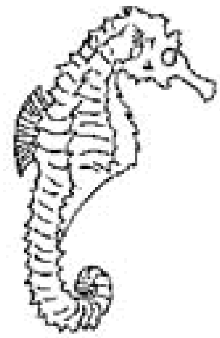
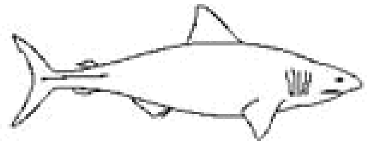


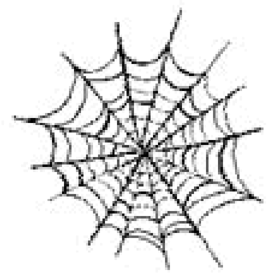
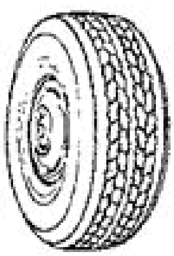
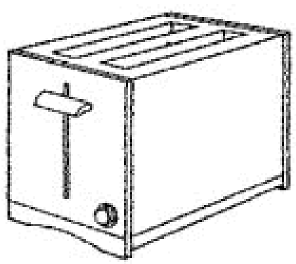


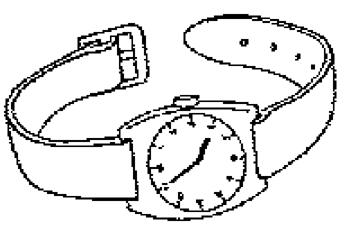






CHUNG-LIN YANG

2100 E. Lingelbach Ln. Apt. 503
Bloomington, IN 47408
U. S. A.
Home: +1 (812) 857 - 3705
Mobile: +1 (812) 360 - 6892
Email: iubmartin@gmail.com

EDUCATION

Ph.D., Linguistics, July 2015
Dissertation: Phonological variation and L2 word learning: The role of orthography in word recognition and production
Minors: Cognitive Science
Speech and Hearing Sciences
Certificate in Cognitive Science - Language and Speech
Indiana University - Bloomington, U. S. A.

M. A., Linguistics, October 2006
Indiana University - Bloomington, U. S. A.

B. A., Russian, June 2003
National Chengchi University, Taipei, Taiwan

RESEARCH INTERESTS

Psycholinguistics
Speech perception and production
Word recognition (both spoken and visual)
Mental lexicon
Cognitive mechanisms underlying second language acquisition

RESEARCH EXPERIENCE

Language and Cognition Lab, Indiana University - Bloomington
Research Assistant to Dr. Chien-Jer Charles Lin, Jan. – Dec. 2012 and Jul. – Dec. 2013

- Assisted in setting up the lab
- Assisted in testing a new electroencephalogram (EEG) system and programming event-related potential (ERP) experiments
- Programmed and assisted in running experiments; analyzed data

Projects:
1. Chinese sentence processing
2. The effect of phonetic orthography (Zhuyin vs. Pinyin) on the perception of Mandarin vowels and tones

Second Language Psycholinguistics Lab and Social Development Lab, Indiana University
Research Assistant to Dr. Nathalie Fontaine, Summer 2013

- Programmed ERP experiments investigating the relationship between children's perception of emotion in speech and child conduct problems.

Second Language Psycholinguistics Lab, Indiana University - Bloomington
Research Assistant to Dr. Isabelle Darcy, Aug. 2008 - June 2010

- Maintained lab facilities and tested equipment
- Assisted in creating experiment materials, programming, running psycholinguistic experiments and data analysis

Projects: 1. Underlying sources of individual differences: cognitive abilities in L2 phonological development
2. Category formation and lexical encoding of new phonological contrasts

Department of East Asian Languages and Cultures, Indiana University - Bloomington
Research Assistant to Dr. Jennifer Liu, Spring 2007
Assisted in collecting data, video-recording Chinese language class meetings, and transcribing the conversations between teachers and students for research purposes.

UNIVERSITY TEACHING EXPERIENCE

Associate Instructor, Dept. of Linguistics, Indiana University, Spring 2015 and Fall 2011
Course title: Introduction to Language and Linguistics (undergraduate level)

Associate Instructor, Dept. of Linguistics, Indiana University, Spring 2014 and Fall 2007
Course title: Language and Religion (undergraduate level)

Associate Instructor, Dept. of Linguistics, Indiana University, Spring 2013
Course title: Introductory Phonetics (graduate level)

Graduate Assistant, Dept. of Linguistics, Indiana University, Spring 2008
Course title: Animal communication (undergraduate level)

GRANTS AND AWARDS

The Householder Research Funds, Department of Linguistics, Indiana University, 2015
(Co-recipient with Kuan-Yi Chao and Yu-Jung Lin)
Travel Grant, Indiana University Linguistics Club, 2015
Conference Travel Award, Department of Linguistics, Indiana University, 2014
Conference Travel Award, Department of Linguistics, Indiana University, 2013
Travel Award, East Asian Studies Center, Indiana University, 2013
Doctoral Grant-in-Aid, Indiana University, 2012
Travel Award, College of Arts and Sciences, Indiana University, 2008.
The Householder Research Funds, Department of Linguistics, Indiana University, 2008
The Householder Research Funds, Department of Linguistics, Indiana University, 2006

PUBLICATIONS

Darcy, I., Park, H., & **Yang, C.-L.** (2015). Individual differences in L2 acquisition of English phonology: The relation between cognitive abilities and phonological processing. *Learning and Individual Differences, 40,* p. 63-72.

RESEARCH PRESENTATIONS

Yang, C.-L. & Darcy, I. (2015, April). *The effect of the orthographic depth of English vowels and consonants on the learning of free variation in an artificial language*. Paper to be presented at the 51st Meeting of the Chicago Linguistic Society. April 23-25, 2015, Chicago, IL.

Hoyniak, C., Bates, J. E., Petersen, I. T., **Yang, C.-L**., Darcy, I., & Fontaine, N. (2015, March). *Atypical Neural Responses to Vocal Fear are Associated with Callous and Unemotional Behaviors in Early Childhood*. Poster presented at the Biennial Meeting of the Society for Research in Child Development. Philadelphia, PA.

Yang, C.-L. & Darcy, I. (2014, October). *When two become one - Orthography helps link two free variants to one lexical entry*. Poster presented at the 168th Meeting of the Acoustical Society of America, Indianapolis, IN.

Yang, C.-L. & Darcy, I. (2014, October). *The effect of orthography on the learning of vocalic and consonantal free variation*. Paper presented at the Second Language Research Forum, Columbia, SC.

Yang, C.-L. & Darcy, I. (2014, October). *Lexical Representation of Free Variants - On the Effect of Orthography on Word Learning*. Poster presented at the 9th International Conference on the Mental Lexicon, Niagara-on-the-lake, Canada.

Yang, C.-L. & Darcy, I. (2013, December). *The role of orthographic information in the learning of allophonic variation*. Paper presented at the 166th Meeting of the Acoustical Society of America, San Francisco, CA.

Lin, Y.-J., **Yang, C.-L**., & Lin, C.-J. (2013, December). *The effect of phonetic orthography on the perception of Mandarin syllables.* Poster presented at the 166th Meeting of the Acoustical Society of America, San Francisco, CA.

Yang, C.-L. & Darcy, I. (2013, November). *The effect of orthography on the mapping of L2 allophonic variants to lexical entries*. Poster presented at Phonology 2013, University of Massachusetts, Amherst, MA.

Yang, C.-L. & Darcy, I. (2013, November). *The effect of orthography on the learning of allophonic variation in L2*. Paper presented at the Second Language Research Forum, Provo, Utah.

Lin, Y.-J., **Yang, C.-L**., & Lin, C.-J. (2013, June). *Syllable perception and the effect of phonetic orthography in Mandarin Chinese*. Poster presented at the 25th North American Conference on Chinese Linguistics at University of Michigan, Ann Arbor, MI.

Darcy, I., Park, H., **Yang, C.-L**., & Gleiser, A. (2012, March). *Individual differences in the development of L2 phonological processing: The contribution of cognitive abilities and executive function*. Poster presented at the 1st Conference on Sources of Individual Linguistic Differences. Ottawa, Canada.

Yang, C.-L. (2011, October). *Vowel undershoot in production of English tense and lax vowels by Mandarin and American speakers.* Poster presented at the 162nd meeting of the Acoustical Society of America, San Diego, California, USA.

Darcy, I., Park, H., & **Yang, C.-L.** (2011, October). *Underlying sources of individual differences: Cognitive abilities in L2 phonological development.* Paper presented at the Second Language Research Forum, Ames, Iowa, USA.

Park, H., Darcy, I., & **Yang, C.-L.** (2011, May). *Underlying sources of individual differences: Cognitive abilities in L2 phonological development.* Poster presented at the 161st meeting of the Acoustical Society of America, Seattle, Washington, USA.

Yang, C.-L. (2009, May). *Effect of word length on vowel production by Mandarin and American speakers: Comparison of /i/ and /ɪ/.* Poster presented at the 157th meeting of the Acoustical Society of America, Portland, Oregon, USA.

Yang, C.-L. (2008, July). *Production of American English /eɪ/ and /ɛ/ by Mandarin speakers: mono- vs. disyllabic words.* Poster presented at the 155th meeting of the Acoustical Society of America, Paris, France.

Yang, C.-L. (2007, May). *The variation between the degree adverbs "chao" and "hen" in Mandarin Chinese spoken in Taiwan*. Paper presented at the 1st Conference on Language, Discourse and Cognition. National Taiwan University, Taipei, Taiwan.

PROFESSIONAL ACTIVITIES

Acoustical Society of America (student member)
Student co-reviewer for Second Language Research