ABSTRACT

Title of Document: THE USE OF SEGMENTATION CUES IN SECOND LANGUAGE LEARNERS OF ENGLISH

Candise Yue Lin, Doctor of Philosophy, 2013

Directed By: Professor Min Wang
Department of Human Development and Quantitative Methodology

This dissertation project examined the influence of language typology on the use of segmentation cues by second language (L2) learners of English. Previous research has shown that native English speakers rely more on sentence context and lexical knowledge than segmental (i.e. phonotactics or acoustic-phonetics) or prosodic cues (e.g., word stresss) in native language (L1) segmentation. However, L2 learners may rely more on segmental and prosodic cues to identify word boundaries in L2 speech since it may require high lexical and syntactic proficiency in order to use lexical cues efficiently. The goal of this dissertation was to provide empirical evidence for the Revised Framework for L2 Segmentation (RFL2) which describes the relative importance of different levels of segmentation cues. Four experiments were carried out to test the hypotheses made by RFL2. Participants consisted of four language groups including native English speakers and L2 learners of English with Mandarin, Korean, or Spanish L1s. Experiment 1 compared the use of stress cues and lexical knowledge while Experiment 2 compared the use of phonotactic cues and lexical knowledge. Experiment 3 compared the use of phonotactic cues and semantic cues while Experiment 4 compared the use of stress cues and sentence context. Results showed that L2 learners rely more on segmental cues than lexical knowledge or semantic cues. L2 learners showed cue interaction in both lexical and sublexical levels whereas native speakers appeared to use the cues independently. In general, L2 learners appeared to have acquired sensitivity to the segmentation cues used in L2, although they still showed difficulty with specific aspects in each cue based on L1 characteristics. The results provided partial support for RFL2 in which L2 learners' use of sublexical cues was influenced by L1 typology. The current dissertation has important pedagogical implication as findings may help identify cues that can facilitate L2 speech segmentation and comprehension.

THE USE OF SEGMENTATION CUES

IN SECOND LANGUAGE LEARNERS OF ENGLISH

Candise Yue Lin

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2013

Advisory Committee:

Professor Min Wang, Chair
Professor Kira Gor
Professor William Idsardi
Professor Rochelle Newman
Professor Meredith Rowe

© Copyright by
Candise Yue Lin
2013

Acknowledgment

Over the past five years, I have received support and encouragement from a great number of individuals. First and foremost, I would like to thank my advisor, Dr. Min Wang, whose expertise, understanding, patience, and generous guidance have made this an intellectually challenging yet rewarding journey. I am greatly indebted to my dissertation committee of Drs Kira Gor, William Idsardi, Rochelle Newman, and Meredith Rowe for their support over the past two years as I moved from an idea to a completed study.

I would like to express my gratitude to Dr. Colin Phillips, Csilla Kajtar and all student and faculty members of the IGERT program. I would not have made it to where I am today without the intellectual and financial support from IGERT. This dissertation research was jointly funded by IGERT and NSF Doctoral Dissertation Research Improvement Grant (BCS-1251589).

I am hugely indebted to Dr. Frank Manis at University of Southern California. As my undergraduate honors thesis advisor, he inspired me to do research in language development. He also provided spaces for my data collection of this dissertation project at USC. I am grateful for my two research assistants, Eva Ortega and Stephanie Li, who helped me with data collection. I would also like to Dr. Scott Jackson who has provided valuable statistical advice. I am indebted to Emily Grossnickle for her help with recording stimuli and proofreading my previous manuscripts.

I would like to thank Chuchu Li, who is a colleague and good friend. She provided much needed encouragement to get me through the ups and downs of graduate school. I am thankful to my two former colleagues in the Bilingual and Biliteracy Research Lab, In Yeong Ko and Say Young Kim, who provided guidance as I first started graduate school five years ago.

Finally, I would like to thank my parents for their unconditional love and support. Most importantly, I want to thank my husband, Chris Lin for the sacrifices he made so I can complete my doctoral education at the University of Maryland, College Park. I am also grateful for Chris for lending his expertise in computer engineering to help program some of the E-prime experiments.

Table of Contents

List of Tables

List of Figures

Chapter 1 - Introduction

When reading texts, locating the beginning and end of a word is simple because there is a visual gap between each word. When listening to speech, however, there is no similarly reliable cue to indicate word boundaries. Imagine a person was watching the news on TV and a reporter said, “morepeoplegettotheirdestinationbycar.” How does the person know that *by* and *car* are two different words instead of one word, *bicar*? This is the problem of segmentation that every language user must solve. One possible solution is to use vocabulary knowledge. A native English listener knows that *by* and *car* are real words in English but *bicar* is not; in such cases, identifying the boundary between *by* and *car* is quick and effortless. However, segmentation becomes more difficult if words in the spoken input are unfamiliar or the listener has a smaller vocabulary size. Imagine a second language (L2) learner of English who does not recognize the word *destination* but knows the word *nation*. The learner may erroneously segment *destine* from *destination*, resulting in an inaccurate interpretation of the spoken phrase. The goal of this dissertation research is to examine the various cues L2 learners rely on in their speech segmentation and establish a Hierarchical Framework that describes the differential weight assigned to the cues as a result of first language (L1) influence. This work has important pedagogical implication as findings can potentially inform educators about how to improve L2 learners’ speech comprehension.

Segmenting continuous speech is a great challenge for both native and nonnative speakers. Speech signal often do not contain breaks at word edges. Even when breaks occur, they do not coincide with perceived word boundaries (Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967). Various cues have been suggested to facilitate

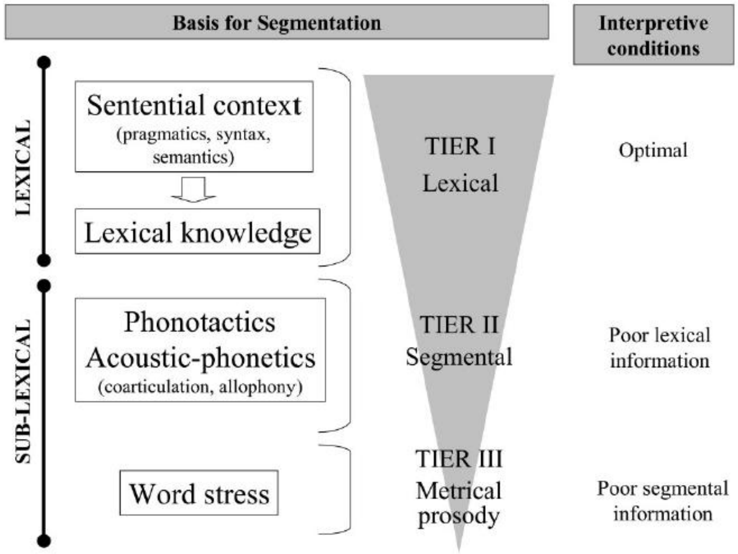

Figure 1. The hierarchical framework of speech segmentation proposed by Mattys et al. (2005).

speech segmentation, including prosodic cues (Cutler & Norris, 1988; Endress & Hauser, 2010), acoustic-phonetic cues (Mattys, 2004; Newman, Sawusch, & Wunnenberg, 2011), lexical cues (Norris, McQueen, Cutler, & Butterfield, 1997), and semantic cues (Mattys, White, & Melhorn, 2005). A Hierarchical Framework (Figure 1) has been proposed to capture the weighted importance of these cues (Mattys, 2004; Mattys et al., 2005). At Tier I of the hierarchy is sentence context which includes semantic, syntactic, and pragmatic cues. Imagine a native English listener trying to recognize the target word *cremate* from the sentence “an alternative to traditional burial is to *cremate* the dead;” although *mate* is a real word, the listener would not predict a word boundary before *mate* because the sentence context is about burial rather than friendship. Also belonging to Tier I in the hierarchy is lexical knowledge. Words that are familiar to the listener can be segmented and identified simply by matching the sound patterns in the signal with the established phonological representation in the lexicon (e.g., match *by car* with /baɪ/ and /kɑː/).

If lexical information is unavailable due to poor listening condition, speakers can rely on segmental cues which are at Tier II of the hierarchy. For example, it has been found that English speakers are more likely to lengthen word-final syllables (Umeda, 1975; Beckman & Edwards, 1987). Speakers who are sensitive to the acoustic cue of duration may predict a word boundary following the lengthened syllable. Another example of a segmental cue is phonotactic probability, which is the likelihood of occurrence of a sound sequence in a certain position in the word (Storkel, 2001). When a native English listener hears the phoneme /f/ (as in *knife)* followed by /m/ (as in *man*), he or she can infer that there is very likely a boundary between /f/ and /m/ since /fm/ is not a legal consonant cluster at the onset position in English. Finally, Tier III of the hierarchy is prosodic cues such as word stress. Stress is lexically contrastive in English and there are minimal pairs of words that differ only in stress location such as *trusty* and *trustee*. However, the location of stress is generally unpredictable in English as stress can fall on any syllable depending on syllable weight and word class. Due to its unpredictability, stress may not provide reliable information about word boundary. In fact, stress cues are only utilized by native English speakers when lexical or segmental information is masked by noise (Mattys et al., 2005). In contrast, stress may be a more useful segmentation cue in languages with demarcative stress such as Hungarian or Finnish.

The Hierarchical Framework is constructed based on findings from native English speakers. It is not clear whether similar weightings of the various segmentation cues can be generalized to L2 learners from different L1 backgrounds. Language typology may result in the assignment of differential weights to the cues proposed in the hierarchy. For example, cues at the lexical tier may be relatively less important for L2 segmentation since L2 learners may need to establish a relatively large L2 lexicon in order to utilize lexical knowledge. L2 learners have also been shown to be less sensitive to pragmatic and syntactic structures in L2 sentences (Bardovi-Harlig & Dornyei, 2012; Jiang, 2007). On the other hand, there is evidence suggesting that L2 learners tend to rely on L1 phonotactic, acoustic, and prosodic cues to segment L2 speech (Altenberg, 2005; Cutler, 2000; Weber, 2000). Thus, it is possible that L1 segmental and prosodic cues may be more important in L2 segmentation. These observations lead to the following research questions which guided the current dissertation project:

1. What is the relative importance of lexical, segmental, and prosodic cues in L2 segmentation?
2. How does L1 typology influence the weighting of these cues in L2 segmentation?
3. Would L2 learners be able to utilize cues that do not exist in L1 to segment L2 speech?

This dissertation research examined English L2 segmentation by learners with Korean, Mandarin, or Spanish L1s. The reason for choosing these languages is that they are typologically different in word stress and phonotactic constraints but share the same property in metrical rhythm which is the durational regularities in speech. Korean, Mandarin, and Spanish are generally classified as syllable-timed languages with roughly equal duration between each syllable whereas English is a stress-timed language with roughly equal duration between each stressed syllable. All three groups of L2 learners share the same disadvantage if they segment L2 speech using L1 metrical cues. However, it is unlikely that they would have the same weightings for other segmentation cues examined in this study (e.g., word stress and phonotactic constraints) due to the typological difference between English and each of the three L1s.

In summary, the goal of this dissertation research is to capture the differential weights given to various segmentation cues based on language typology and propose a model that describes the basis for L2 segmentation. This dissertation is organized as follows: Chapter 2 is a review of the literature on the topic of segmentation by native and nonnative speakers, with a focus on the integration of multiple segmentation cues and the typological differences between L2 English and the L1s of interest. Chapter 3 is the description of the characteristics of the participants who took part in the experiments. Chapters 4-7 are detailed description of the four experiments and their results and discussion. Finally, Chapter 8 discusses the results of the experiments in relation to the proposed revised framework for L2 segmentation and concludes with limitations and directions for future research.

Chapter 2 – Literature Review

Speech comprehension is a daunting task as lexical units or words in continuous speech are not separated by silence. Researchers have identified a number of cues in the speech signal that can facilitate speech segmentation for adult speakers. These cues include sentence context (Dilley, Mattys, & Vinke, 2010; Mattys et al., 2005), lexical knowledge (McQueen, Otake, & Cutler, 2001; Norris et al., 1997; Yip, 2004), acoustics and phonotactics (Davis, Marslen-Wilson, & Gaskell, 2002; Gow & Gordon, 1995; McQueen, 1998; Newman et al., 2011; Shatzman & McQueen, 2006) and prosody (Cutler & Norris, 1988; Cutler, 1997; Endress & Hauser, 2010). Researchers have suggested that speakers do not rely on each of the segmentation cues independently (Mattys 2004; Mattys et al., 2005). Depending on language typology and listening condition, speakers may assign differential weights to the cues and organizd them in a Hierarchical Framework (Mattys et al., 2005).

Compared to the extensive literature in native segmentation, segmentation by nonnative speakers or L2 learners has received relatively less attention. Cutler (2000) has suggested that nonnative speakers tend to apply native segmentation strategies in L2 segmentation. Particularly, previous studies have shown that L2 learners may use L1 prosodic, acoustic, and phonotactic cues to segment L2 speech (Altenberg, 2005; Culter & Otake, 1994; Cutler 1997; Otake, Hatano, & Yoneyama, 1996; Weber, 2000). However, it is unclear whether L2 learners also weigh the segmentation cues differently and integrate them in a hierarchical fashion like native speakers. To the best of the current author's knowledge, only one study has examined the integration of multiple segmentation cues in L2 learners of English (White et al., 2010). Since the perceptual system can capitalize on all relevant information present in the environment (Gomez, 2002; Mattys et al., 2005), it is likely that multiple cues are utilized simultaneously in segmentation. Therefore, research on this topic must examine multiple cues in parallel to reflect this process. The goal of this literature review is to draw from the existing research in segmentation by native and nonnative speakers to propose a model that can make predictions about cue integration in L2 segmentation.

This chapter is divided into three sections. The first two sections review the existing literature in native and nonnative segmentation, respectively. Each section includes the discussion of prosodic, segmental, and lexical cues and their integration. The final section examines the influence of language typology on the use of segmentation cues, particularly focusing on Korean, Mandarin, and Spanish L1s and English L2.

Segmentation in Native Speakers

Humans begin to acquire the ability to understand spoken language before the first day of life. Research has shown that infants as young as two-days old prefer their native language spoken by a stranger over a foreign language spoken by the same person (Moon, Cooper, & Fifer, 1993). Infants' sensitivity to their native language is driven by prenatal auditory experience to their mothers' speech before they were born (DeCasper, & Spence, 1986). At one-month old, infants can discriminate between a pair of minimal contrast, /ba/ and /da/ (Eimas, Siqueland, Jusczyk, & Vigorrito, 1971). By five months, infants can discriminate between languages from different rhythmic classes such as English which is stress-based and Japanese which is mora-based (Nazzi, Jusczyk, & Johnson, 2000). Previous studies have suggested that infants develop the ability to segment continuous speech at around 7-8 months (see Kuhl 2004 for a review). Some of the cues that infants use include distributional probability (Saffran, Aslin, & Newport, 1996), stress pattern (Jusczyk, Houston, & Newsome, 1999), and knowledge about familiar sounds and words (Bortfeld, Morgan, Golinkoff, Rathbun, 2005; Jusczyk & Aslin, 1995). Interestingly, adults, who presumably have extensive knowledge about their native language and acquired full competence in speech comprehension, also rely on some of the same segmentation strategies as infants. In addition, Thiessen and Saffran (2003) found that infants rely more on statistical cue at seven months of age whereas they rely more on stress cues at nine months. Thus, another similarity between infant and adult segmentation is the differential weightings of multiple cues. This section will begin with a review of the use of prosodic, segmental, and lexical cues individually, follow by the discussion of cue interaction and the influence of language typology on cue weightings.

Prosodic Cues

Several prosodic cues have been identified as useful in speech segmentation and they include metrical rhythm (Cutler & Norris, 1988), lexical stress (Endress & Hauser, 2010), and intonation (Diley et al., 2010). While Endress and Hauser argued that stress may be a language-universal segmentation mechanism, Cutler and Norris proposed that metrical segmentation is a language-specific strategy based on metrical rhythm.

Languages differ in their metrical rhythm. Stress-timed languages such as English and Dutch exhibit a strong contrast between strong and weak syllables and strong syllables have longer duration than weak syllables. The Metrical Segmentation Strategy (MSS, Cutler & Norris, 1988) hypothesis predicts that English speakers segment speech at the onset of every strong syllable, which is a syllable with a full vowel that can potentially be stressed. Using the word-spotting paradigm, Cutler and Norris (1988) found that recognition time for a real word, *mint*, is longer in nonsense syllables *mintayve* compared to *mintesh* because *tayve*, a strong syllable with a full vowel, triggers segmentation and lexical search. This delays the recognition of *mint* because the word-final /t/ is initially considered as the onset of the nonword *tayve*. In contrast, the second syllable in *mintesh* is weak and the sequence is not divided, thus there is no obstacle in detecting the embedded word. The strategy of segmenting English speech at strong syllable onsets may be effective for lexical access since approximately 70% of English disyllabic content words begin with a strong syllable (Cutler & Carter, 1987). Dutch has a similar distributional pattern in its vocabulary (van Heuven & Hagman, 1988) and the use of MSS has also been observed in Dutch speakers (Vroomen, van Zon & de Gelder, 1996).

Languages such as French and Spanish have a syllable-based rhythm in which the duration of each syllable is approximately equal. Researchers have found that it was easier for French and Spanish speakers to detect the syllable *ba* in *balance* than in *balcon* Cutler, Mehler, Norris, & Segui, 1986; Sebastian-Galles, Dupoux, Segui, & Mehler, 1992). Speakers of syllable-timed languages tend to place a word boundary between two syllables with approximately equal duration. In this example, the boundary should be placed between *bal* and *con* and *bal* does not match the target *ba*. Japanese is a mora-timed language and each mora is consisted of the syllable nucleus and an optional onset. Thus, Japanese speakers segment the word *pokemon* as po-ke-mo-n. It was found that Japanese speakers segment speech at the boundaries of mora (Otake, Hatano, Cutler, & Mehler 1993). These results suggested that, despite the typological difference in metrical rhythm, native speakers across a variety of languages consistently utilize metrical cues in speech segmentation. One potential problem for the MSS is that there are languages that cannot be strictly classified as syllable-timed or stress-timed. For example, Mandarin Chinese is often classified as a syllable-timed language (Mok, 2008). However, there is a considerable degree of vowel reduction in syllables that carry the neutral tone (Chao, 1968) and vowel reduction shortens the duration of the syllable. It is unclear whether Mandarin speakers segment speech at strong syllable onsets, at syllable onsets, or both.

Endress and Hauser (2010) have suggested that word stress is a less language-specific segmentation cue than metrical rhythm. Stress is mainly realized through three acoustic correlates, namely, intensity, pitch, and duration of the syllable (Hayes, 1995). Languages with more initial-stressed words (e.g., English) tend to rely on pitch to signal stress whereas languages more with final-stressed words (e.g., French) tend to realize stress through duration (Hayes, 1995). In other words, speakers may be able to utilize this regularity in the implementation of initial versus final-stress to locate word boundaries. Endress and Hauser (2010) found that monolingual English speakers were able to identify the target words in French and Hungarian speech. Hungarian has fixed stress on the initial syllable and it belongs to a different language family from French and English. The finding that speakers can use prosody to segment words spoken in entirely unfamiliar languages implies that there is a language-universal mechanism for stress segmentation. However, this conclusion is problematic because Endress and Hauser (2010) did not examine segmentation in languages with no word-level stress such as Korean and Japanese. Speakers of these languages may not be sensitive to the acoustic cues for stress realization. Conversely, English speakers may not use stress cues to segment real Korean or Japanese speech as these languages have neither fixed nor free stress. Since word-stress is not a language-universal prosodic feature, there is no reason to suggest that all speakers can rely on word-level stress cues in segmentation.

Besides prosodic cues at the word level, researchers have also found that speakers utilize phrase-level prosody to disambiguate speech input (Dilley & McAuley, 2008). Speakers tend to group intonation with repeated alternations between a high (H) and a low (L) tone as binaries (e.g., (HL)(HL)(HL) or (LH)(LH)(LH)). Dilley et al. (2010) presented English speakers with four monosyllabic words (e.g., *foot*, *note*, *book*, and *worm*) that can be grouped in more than one way (e.g., *footnote bookworm* or *foot notebook worm*). They were instructed to identify the final word in the phrase. Participants were more likely identify the final word as disyllabic with the HLHL intonation whereas they were more likely to choose the monosyllabic interpretation with the HHLH intonation. Although the researchers did not examine this segmentation cue in other languages beyond English, phrase-level prosody may be less language-specific than word stress. Korean does not have word-level stress, but it has an intonation pattern of high and low tones. Korean speakers can potentially use a similar segmentation strategy like that used by the English speakers in Dilley et al. (2010).

Segmental Cues

Segmental cues that can facilitate native segmentation include acoustic-phonetic cues (Gow & Gordon, 1995; Newman et al., 2011; Shatzman & McQueen, 2006), phonotactic cues (McQueen, 1998; Church, 1983; Vroomen et al., 1998), and coarticulation (Mattys, 2004). There is a high degree of language-specificity for acoustic and phonotactic cues as certain phonemes and rules exist in one but not another language.

In English, allophonic differences of how phonemes are realized in different syllable positions, such as increased aspiration for voiceless stops at word-initial position, can be used by speakers as cues to word boundaries (Christie, 1974; Davis et al., 2002, Nakatani & Dukes, 1977). Lehiste (1960) found that word-initial segments are also longer in duration than equivalent segments that are not word initial. Specifically, lengthening the duration of /s/ delayed word recognition time for Dutch speakers when the target word was *pot* following a word with coda /s/ since longer duration of /s/ activated the competitor word *spot* (Shatzman & McQueen, 2006).

English also has a set of phonotactic constraints that can help a listener identify word boundaries. For example, the phoneme /h/ is always syllable-initial and /ŋ/ is always syllable-final (Church 1983). Native English speakers may predict a word boundary preceding /h/ and a boundary following /ŋ/. Phonotactics can occur at the level of a single sound (i.e. /ŋ/) or at the level of biphone. For example, in English, no /tl/ clusters are allowed within a syllable. Thus, English listeners may predict a word boundary between the two sounds when they hear /t/ followed by /l/. In languages with vowel harmony rules such as Finnish, a clash in vowel harmony often signals word boundary. Vroomen et al. (1998) asked Finnish speakers to identify CVCV target words preceded by a CV prefix and the vowel of the prefix was either harmonious or disharmonious with the vowels of the embedded target. Participants were faster to identify target words preceded by the prefix with a disharmonious vowel. The same result was not found in Dutch or French speakers who do not have vowel harmony rules in their languages. These results suggested that speakers utilize allophonic and phonotactic cues in segmentation, if they are available in the language.

It is important to differentiate between phonotactic constraints that involve absolute legality (the fact that /h/ can only occur syllable-initially) and those that involve probabilities. One example of this is that English content words generally do not end with a lax vowel. English speakers who are sensitive to probabilistic phonotactics may be less likely to place a word boundary after a lax vowel (Newman et al., 2011). However, research has found that the probabilistic phonotactics of a syllable-final vowel were not taken into consideration by English listeners (Newman et al., 2011; Norris, McQueen, Cutler, Butterfield, & Kearns, 2001).

Coarticulation describes a production phenomenon in which segments at the edge of words or phrases tend to have more clear articulation and less overlap with adjacent segments than those within the word (Fougeron & Keating, 1997). In a cross-modal priming task (Mattys, 2004), native English speakers were presented with an auditory phrase consist of a nonword context and the first two syllables of a trisyllabic real word (e.g., *diplo-compro*). The task was to decide whether the letter strings following the auditory phrase represented a real English word (e.g., *compromise*). Segmentation of the target word was made favorable by concatenating the context and the target word (e.g., *diplo-compro*) or made unfavorable by concatenating the first syllable of the target word with the second syllable (e.g., *diplocom-pro*). Concatenation disrupts continuous speech and reduced coarticulation can be a cue for word boundary. Indeed, results showed that lexical decision latency was faster when concatenation coincides with word boundary. Although coarticulation may be a useful cue in English segmentation, it may not be useful for speakers of tonal languages. Xu and Liu (2006) found that in Mandarin, besides the simultaneous onset of the consonant and vowel, there is little or no coarticulation between other adjacent segments. Thus, Mandarin speakers may not use coarticulation as a cue for word boundary in their native language.

Lexical Cues

Lexical cues for segmentation can be divided into cues at the word level and cues at the sentence level. At the word level, Norris et al. (1997) have proposed the Possible Word Constraint (PWC) hypothesis which postulates that the result of segmentation must be a possible word in the listener's language. This constraint operates under the premise that the syllable is the smallest unit that could be a word and all words must contain a vowel. It has been found that the identification of *apple* is easier in *vuffapple* compared to *fapple* because *vuff* is a well-formed syllable but *f* is not (Norris et al., 1997). The use of the PWC as a segmentation cue has been observed in speech segmentation by native speakers of Cantonese (Yip, 2004) and Japanese (McQueen, Otake, Cutler, 2001) (but see Hanuliova, McQueen, & Mitterer, 2010).

Another type of word-level lexical cue is the listener's knowledge about his/her mental lexicon. Mattys et al. (2005) have found that recognition of real words is faster than that of nonwords even though they are embedded in the same phrase context and matched for phonotactic probabilities of the phonemes. For example, the target word *already* is recognized faster in the phrase *animal already* than in *erromal already*. On the other hand, English speakers do not appear to use neighborhood density as a cue for word boundaries. Neighbors are words that differ from one another by the addition, deletion, or substitution of one phoneme in any place of the word. For example, some of the neighbors of *cat* are *mat*, *cap*, *at*, and *cash*. Previous research has shown that spoken words with a dense neighborhood are recognized slower than words in a sparse neighborhood (see Luce & Pisoni, for a review). However, neighborhood density does not seem to influence segmentation. Newman et al. (2011) found that word recognition time did not differ significantly regardless of whether the target words were preceded by syllables with a high or low density neighborhood.

Beyond the word level, there is evidence showing that speakers use semantic context in segmentation. Dilley et al. (2010) asked participants to identify the final word in auditory phrases. The final syllable can be parsed as either a disyllabic or a monosyllabic word (e.g., *turnip* or *nip*). The phrases were either semantically related to the monosyllabic parsing (e.g., *puppy biting cry sister nip*) or to the disyllabic parsing (e.g., *garden veggie crisis turnip*). Participants identified more disyllabic words when the semantic context was consistent with the disyllabic parsing. In another word recognition task, Mattys et al. (2005) found that response latency was faster when the target words were semantically related to the preceding context (e.g., *dressing gown* vs. *mayhem* gown). Sentence predictability also helps speakers to disambiguate sound sequences that can be parsed in more than one way. For example, the spoken word *career* contains the monosyllabic real word *rear*. The MSS predicts that English speakers would take longer to identify *career* because segmentation occurs at every strong syllable onset and the target word would be identified as *rear*. However, participants were significantly faster to identify the visual target of *career* than *rear* when it was preceded by a sentence context consistent with the disyllabic parsing (e.g., *He worked hard for many companies to further his*____). These results suggested that participants may not always rely on metrical rhythm for segmentation if higher-order cues (e.g., semantic) are available.

Cue Integration

Segmentation cues can be largely categorized into three major levels: prosodic, segmental, and lexical. All of the cues may be simultaneously available in the speech input and they may interact to contribute to segmentation jointly. Thus, researchers have recently begun to examine cue integration in native segmentation (Mattys, 2004; Mattys et al., 2005; Newman e al., 2011). Mattys et al. (2005) have proposed the Hierarchical Framework (Figure 1) that captures the weighted importance of the three levels of cues. At the top of the hierarchy is the lexical tier which includes sentence context and lexical knowledge. Tier II is the segmental tier which includes acoustic-phonetic and phonotactic cues. Finally, Tier III is word stress. Speakers only resort to stress cues when speech input is masked by noise which takes away information from the lexical and segmental tiers.

This hierarchy was constructed based on results from a series of six experiments. Experiment 1 compared the use of stress cues and coarticualtion. Results showed that, when speech input was intact, segmentation was faster when concatenation coincided with word boundary regardless of the stress pattern of the words (e.g., initial-stressed vs. medial-stressed). In the noisy condition, speakers segmented initial-stressed words faster regardless of the location of concatenation. Experiment 2 compared the use of stress and phonotactic cues. Segmentation of the target word was made phonotactically favorable or unfavorable by manipulating the probability of the biphone at word boundary (e.g., coda of the preceding word and onset of the target word). For example, the biphone /mk/ has low phonotactic probability and it is more likely to occur across two words than within one word. When speech signal was intact, faster segmentation was observed in phrases with low biphone probability, regardless of stress patterns of the target words. In the noisy condition, segmentation was faster for initial-stressed words regardless of the biphone probability of the phrases.

The use of lexical knowledge and stress cues was compared in Experiment 3. Mattys et al. (2005) found that segmentation was faster when target words were preceded by real words (e.g., *criminal-compromise*) than when they were preceded by nonwords (e.g., *lectinal-compromise*), regardless of the stress patterns of the target words. When noise was added to the auditory stimuli, initial-stressed words were segmented faster than medial-stressed words despite the lexicality of the preceding word. Experiment 4 compared the use of phonotactic and lexical cues. When speech input was intact, speakers showed faster segmentation latency when target words were preceded by real words, regardless of the biphone probability. When lexical information was neutralized by truncating the first syllable of the real word (e.g., *calculus male* → *culus male*), speakers relied more on segmental cues instead.

Experiment 5 compared the use of phonotactic and semantic cues. Word recognition time was faster when target words were semantically related to the preceding words than when they were semantically unrelated, regardless of the biphone probability. Experiment 6 compared the use of stress cues and sentence context. In the intact condition, participants were more likely to parse the target words as iambic words when the sentence context is consistent with the disyllabic meaning (e.g., *contest*). With increasing noise in the speech input, participants were more likely to parse the target words as monosyllabic words (e.g., *test*) despite the sentence context. Results from all six experiments suggested that under optimal listening condition, speakers would consistently rely on semantic and lexical cues. They also assign more weight to segmental cues than to stress cues. However, lower level cues would become useful when the higher level cues are unavailable or inefficient in degraded listening conditions.

Within the segmental level, speakers may assign more weight to allophonic cues than to phonotactic probabilities (Newman et al., 2011). Allophonic cues are the different realization of a phoneme, depending on its position in the syllable. For example, *great eye* and *grey tie* can be differentiated by allophonic cues such as the longer duration of word final /eɪ/ and longer voice onset time (VOT) of the syllable-initial /t/. In this study, phonotactic probability is operationlized as the frequency of vowel occurrence. In English, it is more frequent that syllables end with tense vowels than lax vowels. Newman et al. did not find any vowel effect when there were strong allophonic cues present. This result suggested that there may be another sub-hierarchy within the segmental tier in which English speakers would weight allophonic cues over phonotactic cues.

One major limitation of the Hierarchical Framework is that it is built based on findings solely from native English speakers. It is unlikely that, due to language typology, speakers of other languages also weigh the three levels of segmentation cues in a similar fashion. For example, stress cues in English have relatively small information value about word boundary (Mattys et al., 2005) since the location of stress is unpredictable in English. Stress may be a more reliable segmentation cue if it serves a demarcative function (Jakobson, 1971). In demarcative stress languages such as ancient Hungarian or Finnish, stress is systematically assigned to the *nth* syllable or mora from the word boundary. Speech segmentation can be easily accomplished by applying the same counting strategy to every word. Thus, stress cues may be given the most weight among all segmentation cues by Hungarian speakers. Standard English is traditionally considered to comprise of 24 consonants and most of which can be onsets and codas. The African language, !Xóõ*,* which is spoken in Botswana, has 122 consonants and a very large number of them can only occur word-initially (Trail, 1985). Since the speech signal is processed sequentially from onset to coda, the presence of one of the consonants that only occur word-initially in the spoken input would result in immediate identification of word boundary before the listener hears the coda and completes the word recognition process. In this case, phonotactic cues would probably be given more weight than lexical cues.

Even in languages with less systematic phonotactic cues or with no demarcative stress, there is evidence suggesting that stress cues are given more weight than acoustic-phonetic cues. Vroomen, Tuomainen, and de Gelder (1998) compared word stress and vowel harmony as potential segmentation cues in Finnish. The front-back vowel harmony rule prohibits the co-occurrence of vowels from the front and back harmony class in an uncompounded word. Thus, a clash in vowel harmony in Finnish is typically associated with a word boundary. Finnish has a fixed stress system in which primary stress always falls on the initial syllable (Karlsson, 1999). Stress may be a reliable indicator of word boundaries since it coincides with the beginning of the word. Vroomen et al. (1998, Exp 2) found that Finnish speakers recognized the target words with a stress cue faster than those without a stress cue in the disharmonious vowel condition. In Experiment 3, native Finnish, French, and Dutch speakers were taught an artificial language with vowel harmony rules. There is no vowel harmony in French or Dutch, but the former has fixed stress on the word-final syllable whereas the latter has free stress which predominantly falls on the initial syllable. Finnish speakers only showed the effect of vowel harmony in the absence of stress cues. Dutch speakers did not show any effect of vowel harmony but a robust stress effect. Neither a vowel harmony nor stress effect was observed in French speakers.

In summary, demarcative stress in Finnish, phonotactic cues in !Xóõ, and vowel harmony in Finnish consistently showed that the use of segmentation cues and their weightings can vary greatly depending on language typology. How do humans, with the same innate cognitive system, come up with vastly diversified and mostly language-specific cues to this language-universal problem of speech segmentation? Models constructed based on studies of individual languages cannot provide insight into this inquiry. More comparative studies such as Vroomen et al. (1998) are greatly needed.

Segmentation in Non-Native Speakers

There are four reasons why adult L2 learners would be an ideal population to test whether speakers adjust their weightings of various sources of information depending on their language experiences. First, L2 learners may not be able to use lexical cues and sentence context (e.g., semantics, syntax and pragmatics) efficiently until they have developed a decent size L2 lexicon and constructed semantic representations for L2 words. Thus, semantic and lexical cues may be given less weight in L2 segmentation. Second, adult L2 learners are proficient speakers of their L1 and should have already established weightings for various cues based on the typology of L1. The differential weights given to L1 segmentation cues may have cross-linguistic influence on the process of breaking up continuous speech from a less familiar language. Third, examining multiple groups of L2 learners by using the same set of tasks may reveal the universality of certain cues in L2 segmentation. If L2 learners from a diversity of L1 backgrounds give similar weights to an acoustic cue, presumably this cue is less language-specific and may be utilized by beginning learners with no extensive L2 lexical knowledge. Fourth, identifying cues facilitative to L2 segmentation may inform language instructors how to improve L2 learners' listening proficiency.

Although studies of L2 segmentation are scarce, there is evidence suggesting that learners apply their native strategies to segment L2 speech (Cutler, 2000). Previous research has shown that L2 learners often rely on L1 prosodic, acoustic-phonetic, and phonotactic cues (Altenberg, 2005; Otake et al., 1993; Weber, 2000). Although only one study has examined cue interaction in L2 segmentation (e.g., White et al., 2010), it appears that L2 learners also assign differential weights to cues just like native speakers.

Prosodic Cues

Metrical segmentation is a language-specific strategy and L2 learners may inappropriately apply their L1 metrical segmentation strategy to process L2 speech. For example, French is a syllable-timed language. Yet native English and Dutch speakers, who segment their native speech at every strong syllable onset, do not segment French speech syllabically (Cutler et al., 1986; Cutler, 1997). French speakers segment Japanese and English speech at syllable boundaries even though native Japanese and English speakers do not (Otake et al., 1993; Cutler et al., 1986). Cutler (2000) referred to the application of L1 segmentation strategy to L2 speech as listening to the L2 through the ears of the L1. Unlike native segmentation in which the clarity of the speech signal can change the weights assigned to various levels of cues, it appears that the use of cues in L2 segmentation is dictated by the L1 typology of the L2 learners, not the speech input.

Despite the differences in metrical rhythm, Sanders, Neville, and Woldorff (2002) have shown that Japanese L2 learners of English were able to use stress cues to segment L2 speech. In a phoneme monitoring task, participants were presented with sentences all made up of nonwords but with normal English prosody. Stress pattern was varied by including words that contained target phonemes in different positions and syllables of different stress. The target can be the onset of a stressed or unstressed syllable in the initial or medial position of the word. The Japanese L2 learners showed native-like performance in this task. They were faster to identify the target phoneme when it was in the typical English stress pattern (strong-initial and weak-medial) compared to an infrequent English stress pattern (weak-initial and strong-medial Even though Japanese is mora-based, results from this study suggested that L2 learners were able to adopt a stress-based segmentation strategy in which they tend to segment at the onset of every strong syllable. However, it is unclear whether the L2 learners were indeed segmenting speech at every strong syllable onset or were simply relying on the acoustic cues of the stressed syllables (e.g., higher pitch and intensity and longer duration) to identify the target phonemes. More importantly, using nonwords to construct the sentences took away semantic and lexical cues. Thus, it remains to be seen how stress cues interact with semantic and lexical cues in L2 English segmentation by native Japanese and Spanish speakers.

Segmental Cues

In addition to metrical rhythm, nonnative speakers tend to inappropriately apply L1 phonotactic cues to L2 segmentation. Weber (2000) found that it was easier for English speakers to detect *luck* in *moyshluck* than in *moysluck* because no real English word begins with the sequence *shl-* whereas *sl-* is a legal sequence in English (e.g., *slack*). The opposite result was observed in highly competent German L2 learners of English because *sl-* is a phonotatically illegal onset in German. Altenberg (2005) also found that native Spanish speakers were better at using allophonic cues that exist in their L1 to segment English speech than at using allophonic cues that do not exist in their L1. For example, Spanish has no aspirated voiceless stops (Macpherson, 1975) whereas in English, voiceless stops have stronger aspiration in word-initial position than those in word-final position (e.g., *cat* /kæt/, /k/ has stronger aspiration than /t/) (Christie, 1974). In both Spanish and English, glottal stops can be inserted before word-initial vowels (e.g., *itch* /ɪtʃ/ becomes /ʔɪtʃ/after glottalization) (Borden, Harris, & Rapheal, 2003; Stockwell, Bowen, & Silva-Fuenzalida, 1956).

Spanish speakers were asked to identify a target word in ambiguous English phrases that can be parsed more than one way. Their accuracy was higher when the allophonic cue was glottal stop (e.g., 88% accuracy for the ambiguous phrase *a niche/an itch*) than when the allophonic cue was aspiration (e.g., 55% for *Lou stops/loose tops*). However, their accuracy was the highest when both types of allophonic cues were present (92%). These results showed that Spanish speakers assign more weight to an allophonic cue that is present in their L1, demonstrating the influence of L1 typology on the weighting of segmentation cues. The differential weights assigned to segmentation cues may have important consequences in L2 speech comprehension as the results indicated that Spanish speakers' word identification accuracy was close to chance level for the aspiration stimuli. The absence of aspirated voiceless stops in Spanish created difficulty for utilizing this allophonic cue in L2 segmentation.

Lexical Cues

Only a few studies have examined the use of lexical cues in L2 segmentation. Sanders et al. (2002) have examined the use of lexical and syntactic cues in L2 English segmentation by L1 Japanese and L1 Spanish speakers. Participants listened to English sentences and identified a target phoneme which was located either in the initial or medial syllable. One set of sentences were normal English sentences, which were referred to as *semantic* sentences by the researchers since the sentences provided information about meaning. Another set of sentences were *syntactic* sentences as all open-class words in the sentences were replaced by nonwords. These sentences had intact syntactic structure but they were no longer meaningful. Both groups of L2 speakers and native English speakers showed higher phoneme identification accuracy with the semantic sentences than with the syntactic sentences. This result suggested that L2 learners were able to utilize sentence context in segmentation, regardless of their L1 backgrounds.

Norris, McQueen, Cutler, Butterfield, and Kearns (2001) showed that the Possible Word Constraint can be used in nonnative segmentation even though the constraints for what constitutes a word differ across languages. For example, an open syllable with a short full vowel such as /ɛ/ or /æ/ cannot be a word in English whereas open syllables with short vowels are acceptable words in French or Japanese (Cutler, 2001). Native English speakers found it as easy to detect words (e.g., *perturb /* pərˈtɜrb/) in contexts consisting of syllables which cannot be English words (e.g., /dɛ/ in /dɛpərˈtɜrb/) as in contexts consisting of syllables which can be English words (e.g., /dɑ/ in /dɑpərˈtɜrb/) (Norris, et al., 2001). Only single consonant contexts (e.g., *f* in *fapple*) appear to make word detection difficult and this is a consistent finding across languages (Cutler, 2001). Thus, the use of lexical cues in L2 segmentation may not be influenced by L1 typology.

The use of lexical cues in L2 segmentation may be influenced by L2 proficiency rather than L1 typology. It has been found that the link between the lexical representation of L2 words and the corresponding conceptual representation is weaker compared to that between L1 words and concepts (Kroll & Stewart, 1994). With increased proficiency, the direct link between L2 lexical representation and the corresponding conceptual representation will be strengthened. Thus, L2 learners may have to achieve a certain level of L2 proficiency in order to take advantage of the L2 lexical and semantic cues.

Cue Integration

To the best of the current author's knowledge, only one study has examined cue interaction in L2 English segmentation. White et al. (2010) compared the use of stress and lexical cues in native Hungarian speakers who have achieved various proficiency levels in English. Participants were divided into two groups, beginning and intermediate learners, to examine whether L2 learners' use of lexical cues differed by their level of L2 proficiency. Unlike English, Hungarian is a fixed stress language in which stress placement is always word-initial. The predictability of stress location in Hungarian may render stress a more reliable segmentation cue in the participants' L1. Therefore, the second research question in this study is whether the presumed heavier weight of stress cues in L1 Hungarian segmentation would transfer when Hungarian speakers segment English speech.

White et al. (2010) utilized the cross-modal priming paradigm. In each trial, participants were asked to listen to a five-syllable phrase (e.g., *anythingcorri*) with visual presentation of a three-syllable letter string (e.g., *corridor*) 100ms after the offset of the auditory prime. The participants' task was to determine whether the visual stimulus was a real English word. The first three syllables in the auditory phrase were referred to as the *context* (e.g., *anything*) while the last two syllables were referred to as the *prime* (e.g., *corri*). The independent variables were the lexicality of the context (e.g., *anything* or *imoshing*) and the stress pattern of the prime (e.g., *corri*- or *confu*-).

Results showed that both the native English speakers and the Hungarian L2 learners responded faster to target words following real word contexts than nonword contexts. The magnitude of priming did not significantly differ between initial-stressed and medial-stressed words regardless of the lexicality of contexts. These results suggested that both native and nonnative speakers used lexical knowledge in segmentation. The absence of any advantage in initial-stressed prime also suggested that neither of the two groups use metrical segmentation as predicted by Cutler and Norris' MSS (1988). Furthermore, L2 speakers were divided into four groups based on their performance on a proficiency test. The researchers did not find a lexical priming effect in the lowest proficiency group, suggesting that L2 speakers with a small vocabulary size did not utilize lexical knowledge in L2 segmentation.

White et al. (2010) concluded that the Hierarchical Framework can also be generalized to Hungarian L2 learners who were able to exploit lexical cues in segmentation. The absence of any stress effect is also consistent with the hierarchy which predicts that speakers only resort to prosodic cues in degraded listening condition. However, Hungarian learners may not use stress cues in segmentation not because they are relatively less important than lexical cues but because Hungarian speakers do not encode stress in their phonological representation. Previous research has shown that speakers of fixed stress languages, such as French and Turkish, cannot discriminate minimal stress pairs if the task prevents them from using acoustic cues (e.g., pitch, intensity, and duration) by imposing a high demand on working memory (Dupoux, Peperkamp, &, Sebastian-Galles, 2001). Thus, it is possible that Hungarian learners of English could not utilize stress cues because they do not encode stress in phonological memory. Peperkamp (2004) has suggested that for fixed-stressed languages like Hungarian, stress can be assigned postlexically whereas for free-stressed languages like English, stress should be encoded in lexical representation due to the unpredictability of stress location.

Language Typology of Korean, Mandarin, and Spanish and the Revised Framework for L2 Segmentation

The literature reviewed so far suggests that language typology greatly influences the use of segmentation cues. For example, Finnish speakers use the clash of vowel harmony as a cue for word boundary but not French or Dutch speakers because vowel harmony does not exist in French or Dutch (Vroomen et al., 1998). In L2 segmentation, French speakers segment Japanese speech syllabically even though Japanese is a mora-based language (Otake et al., 1993). This dissertation project will examine L2 segmentation by learners with Korean, Mandarin, or Spanish L1 and explore how this language typology influences the interaction of segmentation cues. The reason for choosing these three languages is that they are similar to English in one aspect or another. For example, in terms of phonotactics, Mandarin, Korean and English have a phoneme /ŋ/ that can only occur syllable-finally but this phoneme does not exist in Spanish. On the other hand, /s/ is a legal coda in both English and Spanish but not in Mandarin and Korean. Korean and Mandarin speakers may be better at segmenting words that end with /ŋ/ whereas Spanish speakers may be better with words that end with /s/.

Prosodic Cues

In terms of prosody, Korean is typologically farthest from English compared to Mandarin and Spanish. Korean does not have lexically contrastive stress (Jun, 2005; Sohn 1999), and there is no minimal word pair differing in stress alone. Also, Korean does not have fixed stress at the word level (Jun, 1995). A previous study conducted by the current author found that Korean speakers have difficulty discriminating minimal stress pairs (e.g., /'mipa/ and /mi'pa/) and they do not encode stress in phonological representation (Lin, Wang, Idsardi, & Xu, under review). Stress cues may not be used by Korean speakers to segment English speech and can be absent from Korean speakers' hierarchy. On the other hand, both Mandarin and Spanish have lexically contrastive stress and speakers of both languages have stress representation in phonological memory (e.g., Dupoux et al., 2001; Lin et al., in preparation). Mandarin may be more similar to English, compared to Spanish, for two reasons. First, Spanish does not have vowel reduction whereas a change in vowel quality is one of the cues signaling stress in Mandarin and

English (Fry 1958; Shen, 1993). Second, stress mostly falls on the penultimate syllable in Spanish (Navarro, 1966) whereas the predominant stress pattern is initial-stress in English (Cutler & Carter 1987). In Mandarin, a weak syllable cannot be word-initial (Duanmu, 2007) and thus initial-stress is a more frequent pattern. Since penultimate stress does not coincide with the beginning or the end of the word, Spanish speakers may give less weight to stress cues in segmentation. In contrast, Mandarin speakers can utilize stress cue in segmentation since a weak syllable would not signal word onset. Overall, stress cues may be given more weight than lexical cues because high L2 lexical proficiency is not necessary to utilize stress cues if the L2 speakers can rely on their sensitivity to stress as a result of exposure to a L1 with contrastive stress.

Segmental Cues

Previous research has demonstrated that the most reliable cues to word boundaries are located at word onset (Davis et al., 2002). Cross-linguistically, onsets seem to be marked by features such as aspiration (Lehiste, 1960; Nakatani & Dukes, 1977), lengthening of word-initial phonemes and syllables (Klatt, 1973; Gow & Gordon, 1995, Quene, 1992), and laryngalization and glottalization of word-initial vowels (Lehiste, 1960; Nakatani & Dukes, 1977). For example, in English, word-initial voiceless stops /p t k/ have longer VOT and are aspirated and word-final voiceless stops have shorter VOT and are often unaspirated. In Mandarin, the distinction between aspirated and unaspirated is not allophonic, but phonemic. For example, /tā/ means *to take some form of transportation* in English whereas /t^hā/ means *he/she*. The unaspirated stops can become voiced /b d g/ before an unstressed vowel (Duanmu, 2007). In Korean, there are three types of stops: aspirated (e.g., 풀 “pul” *grass* or *glue*), tense (e.g., 뿔 “ppul” *horn*) and lax (e.g., 불 “bul” *fire* or *light*) (Sohn, 1999). Aspirated stops can only occur syllable-initially in both Mandarin and Korean. On the other hand, Spanish does not have aspirated voiceless stops (Macpherson, 1975). Hearing aspirated /p t k/ in English maybe a cue for word initiality for Mandarin and Korean speakers but not for Spanish speakers. Indeed, Altenberg (2005) found that Spanish speakers were significantly worse than native English speakers at segmenting speech using aspiration cues. However, Spanish speakers can utilize the glottalization of word-initial vowels (e.g., *itch* /ʔɪtʃ/) as a cue for word boundaries in L2 English since this allophonic difference occurs in both Spanish and English (Altenberg, 2005).

In terms of phonotactics, Mandarin is typologically farthest from English. Mandarin has 19 consonants and only two of them, /n/ and /ŋ/, are allowed in the coda position (Duanmu, 2007). When hearing a vowel followed by a consonant that is not /n/ or /ŋ/, a native Mandarin listener is likely to place a word boundary after the vowel. Korean has a less restricted set of codas. Seven of the 22 consonants in Korean can occur syllable-finally and these include /p, t, k, m, n, ŋ, l/ (Sohn, 1999). Compared to Mandarin, coda cues may be less informative for word boundary in Korean. Both Korean and Mandarin do not allow consonant clusters at the onset position, although a glide (/j, w/) may follow a consonant in onset. Spanish does not allow consonant clusters in syllable coda except in loanwords (Dalbor, 1997) and no /sC-/ cluster is allowed in syllable onset (Stockwell & Bowen, 1965). Thus, hearing a consonant cluster is a strong cue for word boundary for all three groups of L2 learners.

Comparing across the four languages, Korean, Mandarin and English have the same phonotactic rule for /ŋ/ which limits the occurrence of this phoneme to syllable-final position. In Spanish, /ŋ/is not a phoneme but an allophone of /n/, occurring only before /g k h/ (MacPherson, 1975). Previous research has shown that speakers do not perceive a phoneme and its allophone as two distinct sound categories (Kazanina, Phillips, & Idsardi, 2006). Thus, /ŋ/ can be a strong cue for word boundary for Korean, Mandarin, and English speakers but not for Spanish speakers since Spanish speakers would perceive /ŋ/ as /n/. On the other hand, /s/ cannot occur word-finally in Korean and Mandarin, speakers from these L1 groups may erroneously consider /s/ as the onset of the following word.

Although /n/ can occur both word-initially and finally in all four languages, the degree and direction of nasalization in coarticulation may vary. American English shows extensive anticipatory (i.e. right-to-left) vowel nasalization (Clumeck, 1976; Krakow, 1989, 1999). In an acoustic analysis of the speech produced by four English speakers from Michigan, Tanowitz and Beddor (1997) found that 80% of the vowels in CVN(C) syllables are nasalized. Carryover (i.e. left-to-right) vowel nasalization in NVC words has also been documented in American English (Sole, 1992). In addition, word-final /n/ may assimilate to the place of articulation of a following word-initial consonant (Local, 2003). For example, in *ran quickly*, /n/ may be realized as /ŋ/. However, there is evidence suggesting that such assimilation to a velar nasal is not identical with forms such as *rang* which has final citation-form velars (Kelly & Local, 1989). The anticipatory and carryover nasalization as well as the assimilation rule for coda nasals are present in many dialects of Spanish (Boomershire, 2006; Sole, 1992). Particularly, Mexican and Caribbean Spanish have a word-final neutralization rule which results word-final alveolar nasals preceding a pause or a vowel (Boomershie, 2006; Pineros, 2006). Since the phoneme /ŋ/ does not exist in Spanish, it is not clear whether Spanish speakers would consider velarized /n/ in Spanish and /ŋ/ in American English as identical.

Korean has a number of assimilation rules associated with nasalization (Davis & Shin, 1999): 1) a stop nasalizes before a nasal /sip-nyən/ → [sim.nyən]; 2) /n/ becomes a lateral when immediately before a lateral /non-li/ → [nol.li]; 3) /l/ becomes nasalized when after a non-coronal nasal /kam-li/ → [kam.ni]; and 4) nasalization of obstruent-liquid sequences /pəp-li/ → [pəm.ni]. In standard Mandarin spoken in Taiwan, the syllable-final distinction between /n/ and /ŋ/ tends to be dropped whereas the distinction is maintained in standard Mandarin spoken in Mainland China (Lin, 2002). In light of this observation, the current study ensured that participants in the Mandarin group are from mainland China. All four language groups display a high degree of variability of nasality in word-initial and word-final positions. However, this is unlikely to affect listeners' sensitivity to the phonotactic constraints associated with /n/ and /ŋ/.

Korean and Mandarin speakers are more likely to associate word-final /n/ with the previous string compared to word-final /s/ since /s/ is not a legal coda in their L1. Thus, they are more likely to associate word-final /s/ in English words with the next string. In comparison, /n/ and /s/ are both legal onsets and codas in English and Spanish. Based on the phonotactic probability calculator (Vitevitch & Luce, 2004), word-initial /s/ has a position-specific probability of .1024 while word-final /s/ has a position-specific probability of .0101. Word-initial /n/ has a position-specific probability of .0238 whereas word-final /n/ has a position-specific probability of .0583. If English speakers are sensitive to probabilistic phonotactics, they may be more likely to associate /s/ with the following string since word-initial /s/ has a higher likelihood of occurrence than word-final /s/. In contrast, English speakers may be more likely to associate /n/ with the previous string as word-final /n/ has higher position-specific probability than word-initial /n/. It should be noted that the phonotactic probabilities reported here are calculated based only on single morpheme entries. If morphologically complex words are included in the calculation, it is very likely that word-final /s/ has a higher likelihood of occurrence than word-final /n/ since –s is the inflectional morpheme for plurality in nouns and verbal agreement for third-person singularity. Dewey (1950) analyzed the relative frequency of English speech sounds in 100,000 words selected from a variety of sources including newspaper, speech, and correspondence. Multimorphemic words such as *knows*, *beginning*, and *organization* were included in the materials. Dewey found that word-initial /n/ occurs in 2,590 words (2.59% out of 100,000) whereas word-final /n/ occurs in 8,740 words (8.74% out of 100,000). In comparison, word-initial /s/ occurs in 5,575 (5.575% out of 100,000) while word-final /s/ occurs in 4,630 (4.63% out of 100,000). Listeners who are sensitive to these phonotactic probabilities may experience more difficulty segmenting words with coda /s/ as the likelihood of /s/ being word-initial or word-final is similar. In contrast, segmentation of words with coda /n/ may be easier since /n/ has a higher frequency in word-final than in word-initial position.

The phonotactic probability calculator (Vitevitch & Luce, 2004) is not available in Spanish, it is not clear whether Spanish speakers are more likely to associate /n/ or /s/ with the preceding or following string. Considering most Spanish words end in a vowel (approximately 73%) (Guion, Harada, & Clark, 2004), it is possible that both /n/ and /s/ have low phonotactic probabilities in word-final position. However, word-final /s/ has more phonological variation in Spanish than word-final /n/ and this may result in increased difficulty for the segmentation of words with /s/ coda compared to those with /n/ coda. Word-final /s/ can undergo deletion, aspiration, voicing, or can be retained as /s/ based on the particular dialects of Spanish (Boomershine, 2006). For example, the processes of deletion and aspiration do not occur in Mexican Spanish but they do occur in Puerto Rican Spanish. On the other hand, there is a word-final neutralization rule for coda nasals in all dialects of Spanish (Boomershine, 2006). This neutralization rule results in word-final alveolar nasals preceding a pause or a vowel. The coda nasal assimilates the place specification of the following consonant. In both a lexical decision task and a naming task of Spanish words, Bommershine (2006) found that both native speakers of Puerto Rican Spanish and native speakers of Mexican Spanish were significantly slower to respond to words with coda /s/ than those with coda /n/. It appears that the phonological variation of word-final /s/ has a greater impact on the processing of L1 Spanish words compared to word-final /n/.

Overall, both stress and segmental cues can be categorized as typological cues because the learners' weighting of these cues in L2 segmentation may be influenced by L1 typology. Korean learners may only rely on segmental cues since they do not have a phonological representation of stress and this may result in their inability to use word stress as a cue for word boundaries. On the other hand, stress is contrastive in both Mandarin and Spanish, stress cues may be useful in L2 segmentation for Mandarin and Spanish speakers. However, Spanish speakers may assign more weight to segmental cues than stress cues since their sensitivity to the penultimate stress may not be helpful in segmenting English speech. Although the stress patterns in Mandarin are similar to those in English, Mandarin speakers may still weigh segmental cues over stress cues. Mandarin speakers may rely more on acoustic-phonetic cues for word boundary if they encounter words with unstressed initial-syllables. In addition, the presence of /ŋ/ or aspiration is a strong cue for word ending or word onset, respectively.

Lexical Cues

In contrast to stress or phonotactic cues, there may be smaller language group differences in the use of lexical and semantic cues if English proficiency is matched across the L2 learners. The use of lexical cues involves learners' knowledge and language skills in L2. Learners who are more proficient tend to have a larger and more integrated L2 lexicon. There is no need for the L2 listener to apply native segmentation strategy if he/she readily recognizes the word. Thus, for learners with advanced L2 proficiency, their weightings of cues may be similar to those of native speakers. However, for learners who are still developing their L2 lexicon, they may rely more on sublexical cues than lexical cues, regardless of their L1 backgrounds. Similarly, the use of semantic cue may also require L2 learners to establish a direct link between L2 lexical representations and conceptual representations. Learners with low or intermediate L2 proficiency have to go through L1 translation equivalents to activate the conceptual representation for the target L2 words (Kroll & Stewart, 1994). If all three groups of L2 learners are at a similar stage where they are establishing connections between L2 words and concepts, they should all assign less weight to semantic cues in segmentation since they cannot access L2 semantic representation efficiently.

Within the lexical tier, L2 learners may assign more weight to word-level cues (e.g., lexical knowledge) than sentence level cues as the formal representation of a word is learned much earlier than its semantic and syntactic representations (Jiang, 2002).

Elgort (2011) has demonstrated that L2 learners of English can quickly establish lexical entries for newly learned L2 vocabularies in the mental lexicon. In Elgort's study, advanced L2 learners were asked to use only flash cards to learn pseudowords created by changing one letter in real words (e.g., INFECENT obtained from *indecent*). After one week of learning, participants were tested in three priming experiments including form (e.g., *bunction* – FUNCTION), repetition (e.g., *obsolate* – OBSOLATE), and semantic priming (e.g., *veranda* – BALCONY). All three experiments showed reliable priming effects with the pseudowords as primes. These results suggested that learners with advanced L2 proficiency may develop lexical knowledge quickly. On the other hand, it has been found that compared to native English readers, L2 learners with Mandarin L1 did not show any delay in reading ungrammatical sentences that lack the plural –*s* marker since plurality is not marked by an inflectional morpheme in Mandarin (Jiang, 2007). Similarly, Marinis, Roberts, Felser, and Clahsen (2005) have shown that English learners with Chinese, Japanese, German, or Greek L1 all displayed nonnative like processing mechanisms for English sentences involving *wh*-movement even though Chinese and Japanese do not have *wh*-movement whereas German and Greek do. These results suggested that the lack of sensitivity to syntactic cues is not influenced by typological distance between the L1 and L2 syntactic structures. Thus, it is likely that lexical knowledge is a more accessible segmentation cue for L2 learners than semantic or syntactic cues.

The Proposed Revised Hierarchical Framework for L2 Segmentation

Stress and lexical cues were mostly investigated in native segmentation or in monolingual speakers processing various foreign languages. To the best of the current authors' knowledge, no previous study has examined segmentation in speakers from multiple L1 groups processing one common L2. One can only claim universality for a segmentation cue when L2 learners from two or more different L1 backgrounds exhibit the same consistent utilization of this cue. Examining multiple groups of L2 learners would also test the generalizability of a model that describes the use of segmentation cues in L2 speech comprehension. Since weights are likely to be given to cues differentially based on language typology, results from this dissertation project would motivate the construction of a Revised Framework of L2 Segmentation (RFL2, Figure 2) that captures the relative weight of various segmentation cues for English as a L2.

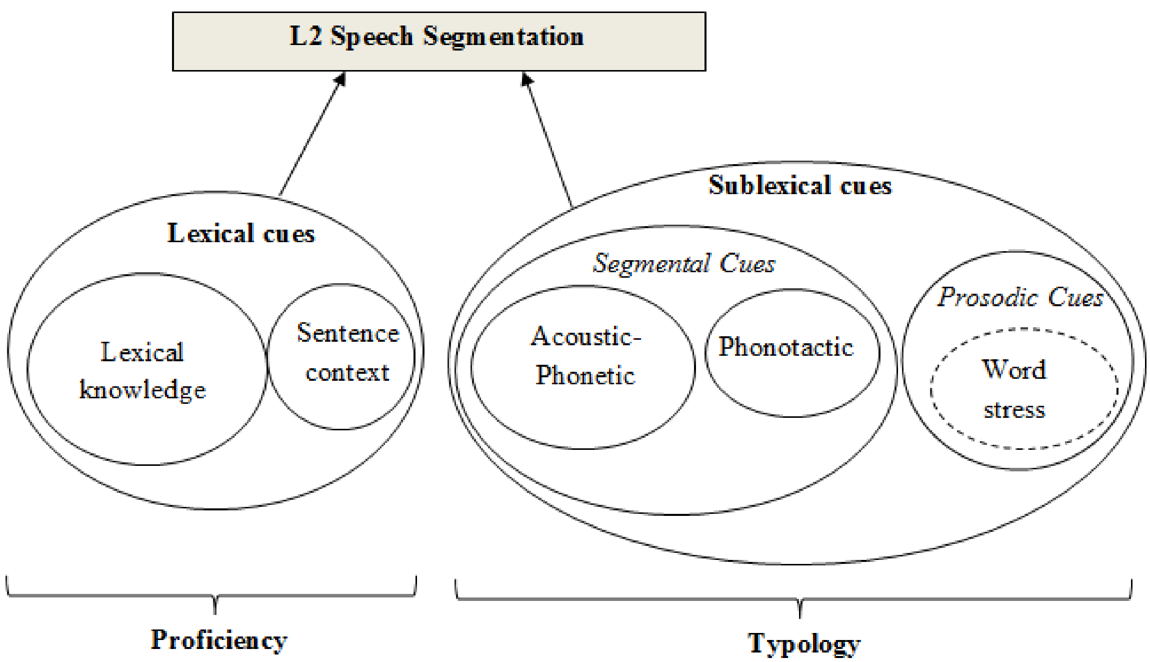

Figure 2 The proposed revised framework for L2 speech segmentation.

The RFL2 makes six predictions: 1) Learners with low to intermediate L2 proficiency will weigh sublexical cues over lexical cues; 2) Among sublexical cues, segmental cues will be relatively more important than prosodic cues for L2 learners; 3)

Learners whose L1 does not have word-level stress will not be able to utilize stress cues to segment an L2 with word stress; 4) Among lexical cues, learners will weigh lexical knowledge over semantic/sentence context; 5) Within the segmental level, acoustic-phonetic cues outweigh phonotactic cues.

In this model, both sublexical and lexical cues play important roles in L2 speech segmentation, which is indicated by the arrows pointing upward. The size of the circle represents the relative importance of the corresponding cues. It is hypothesized that less advanced L2 learners will rely on sublexical cues more than lexical cues since the efficient use of both lexical knowledge and sentence context may require the learner to have a firmly established L2 lexicon, direct links between L2 lexical and semantic representations, and automatized sensitivity to L2 syntactic and pragmatic structures. Similar to Mattys et al.'s (2005) Hierarchical Framework for native English speakers, it is predicted that segmental cues will outweigh prosodic cues for L2 learners. Since previous research has not compared the relative weights of segmental and prosodic cues in L2 segmentation, our hypothesis is based on findings from L1 studies. It is hypothesized that learners are less likely to use stress cues if their L1 does not have word-level stress. The circle with dashed line represents the possible absence of stress at the level of prosodic cues.

Within the lexical level, it is hypothesized that L2 learners will weigh lexical knowledge over sentence context as the formal representation of a word is learned much earlier than its semantic and syntactic representations (Jiang, 2002). Thus, learners may be able to identify a lexical item as a real word but not be able to understand its meaning or use it in a sentence. Finally, it is hypothesized that acoustic-phonetic cues will be slightly more important than phonotactic cues. Newman et al. (2011) found that native English listeners used allophonic cues (e.g., longer voice onset time of syllable-initial /p, t, k/) but not phonotactic probabilities (English syllables are more likely to end with tense than lax vowels, e.g., /i/ vs. /I/) when both cues were present. Since no previous research has compared the use of these two cues in L2 segmentation, our hypothesis is based mainly on L1 research. However, testing this explicitly across a variety of L1 listeners would allow us to examine the generality of the previous L1 findings.

The proposed model also postulates that when L2 learners use lexical cues, they are essentially relying on their L2 proficiency. The knowledge of what constitutes a real word in English and sensitivity to the semantics and syntax in sentence context are developed through the process of L2 acquisition. Larger vocabulary size, more established L2 lexicon, and more efficient semantic and syntactic processing in sentence comprehension are all indications of improved L2 proficiency. Better knowledge of the lexicon and higher sensitivity to sentence context are hypothesized to have a positive effect on L2 speech segmentation. In contrast, cues at the sublexical level are typological cues because the weighting varies by the particular pattern of segments and suprasegments in different languages. For example, word stress is relatively less important segmentation cue in English (Mattys et al., 2005) since stress location in English is less predictable than that in languages such as Hungarian or Finnish in which stress location is fixed. The relative weightings of sublexical cues may vary by L2 typology (if the learners have acquired complete control of the cue cues used in L2) or transfer from L1 phonological structures.

Note that the RFL2 does not make prediction about the time course of cue arrival. The RFL2 describes the relative importance of multiple levels of cues when they are simultaneously available in the speech signal. The hierarchical order of cues is not fixed. Rather, the weighing of cues can change as a result of increasing L2 proficiency, the noise-level of the listening condition or the timing of cue arrival. Advanced L2 learners' hierarchy may be similar to that of native speakers. When the speech signal is degraded, L2 learners may rely more on prosodic cues than segmental cues. In natural conversation, it is very likely that some cues would be available in the speech signal sooner than other cues. For example, prosodic and segmental cues may be available in every word. At the beginning of the sentence, when semantic context or syntactic structure is not entirely clear to the listener, the sublexical cues may play a more important role than the lexical cues. However, as more information about sentential context becomes available, the lexical cues may outweigh the sublexical ones. In addition, the cues that arrive earlier may have an inhibitory effect on the cues that arrive later. Speakers have to revise their early hypotheses about where a word boundary falls as an utterance is unfolding and nonnative speakers have been shown to be more reluctant than native speakers to revise their initial interpretation (Field, 2008). Thus, it is possible that L2 learners may rely more on prosodic and segmental cues than lexical cues since the former are available earlier. The interaction among different levels of cues is dynamic and fluid and it is influenced by many aspects of language processing. The RFL2 reflects the influence of one of these aspects, namely, L1 typology.

The goal of this dissertation project was to examine the differential weights given to segmentation cues by L2 learners of English from three L1 backgrounds. L2 proficiency was taken into account in statistical analyses to ensure any difference in the weighting of cues cannot be attributed to variation in L2 attainment. Four experiments were designed and each experiment compared a pair of cues, one from the lexical level and one from the sublexical level. Experiment 1 compared the use of stress cues and lexical knowledge. Experiment 2 compared the use of phonotactic cues and lexical knowledge. Experiment 3 compared the use of phonotactic cues and semantic cues. Finally, Experiment 4 compared the use of stress cues and sentence context cues. Prediction 1 of RFL2 would be tested in all four experiments while Prediction 3 would be tested in Experiments in 1 and 4 by examining Korean speakers' use of stress cues. Prediction 4 would be tested by comparing the data between Experiments 1 and 4 and between Experiments 2 and 3. The influence of L1 typology on cue weighting at the sublexical level would be tested in all four experiments by examining L2 learners' use of cues that are congruent or incongruent with the properties in L1 phonology.

Summary

In the limited number of L2 segmentation studies, cues have often been investigated individually. Since the perceptual system can capitalize on all relevant information present in the environment (Gomez, 2002; Mattys et al., 2005), it is likely that multiple cues are integrated and utilized simultaneously in segmentation. Research on this topic must examine multiple cues in parallel to reflect this process. In addition, the influence of language typology on the use of segmentation cues has rarely been considered. Oftentimes researchers compare the performance of native speakers with L2 learners from only one L1 background and make the conclusion that L2 learners have difficulty acquiring a certain feature in the L2. However, it is not clear whether this difficulty stems from low L2 proficiency or typological differences between the particular L1-L2 pair. Instead, a fair comparison is to analyze performances from multiple groups of L2 learners with different L1 backgrounds but matched in L2 proficiency. The three L2 groups examined in this dissertation present a unique case of testing the influence of language typology on weighting segmentation cues. Findings from this research will provide empirical evidence for the RFL2 (Figure 2) that is parallel to the monolingual English model proposed by Mattys et al. (2005, Figure 1). The languages examined in this dissertation belong to different families (Korean: Altaic, Mandarin: Sino-Tibetan, Spanish: Romance, and English: Germanic) and this diversity increases the generalizability of the findings. The RFL2, if supported, will be a significant contribution to the literature as it may reliably describe the differential weights given to segmentation cues by L2 learners of English with various L1 backgrounds.

Chapter 3 – Participants and Proficiency Measures

There were four groups of participants, a monolingual English group and three L2 learner groups with L1 Mandarin, Korean, or Spanish. A total of 131 participants were tested, including 39 Mandarin speakers, 34 Korean speakers, 23 Spanish speakers, and 35 English speakers. Participants completed four experiments in two separate sessions, with an average one-week interval between each session. The Language Experience and Proficiency Questionnaire (LEAP-Q, Marian, Blumenfeld, & Kaushanskaya, 2007), the cloze test, and Experiments 1 were administered in the first session and Experiments 2, 3, and 4 were administered in the second session. The order of tasks within each session was counterbalanced across participants.

All participants in the Mandarin group finished the tasks in both of the two testing sessions. Three participants in the Korean group and one participant in the Spanish group did not return for the second session. Thus, these four participants only had data for Experiment 1 and the proficiency measures. Although people from mainland China and Taiwan speak Mandarin, actual production varies widely depending on region. Particularly, light syllables (equivalent of unstressed syllables in English) rarely occur in the vocabulary of Mandarin speakers from the south (i.e. Canton and Hokkien). In addition, many residents of the Southern provinces speak a language or regional dialect that is mutually incomprehensible with Mandarin as their L1 (i.e. Shanghainese, Cantonese, and Taiwanese) and learn Mandarin as L2 in school. To ensure that participants in the Mandarin group indeed speak standard Mandarin (based on the Beijing dialect) as their L1, they were asked to report their city of origin and familiarity with a regional dialect or language besides Mandarin. Nine participants were excluded from the

Mandarin group due to their extensive experiences with a regional dialect or language other than Mandarin.

All participants completed the LEAP-Q. In this questionnaire, participants provided information about age of acquisition (AOA), length of U.S. residence, and amount of language use in L1 and L2. Participants also rated their speaking, listening, and reading abilities in L1 and L2. Another proficiency measure was the cloze test (Bachman, 1982), a fill-in-the-blank activity that assesses English syntactic and lexical proficiency. This test was not timed and the total possible score is 50. Table 1 shows the demographics of the four language groups.

All participants were current students, visiting scholars, or had at least a bachelor degree. Participants used English regularly as they are working or studying in a higher-education institution in the United States. All three L2 groups reported having 40% or higher in the amount of L2 usage (Table 1). All participants in the L2 groups may be considered as early L2 learners as their average age of English acquisition is at age 10 or younger. In their native countries, most L2 learners might have acquired English through formal classroom instructions. When they moved to the U.S., learning might take the form of natural exposure to native speakers. Thus, length of U.S. residence indicates the length of immersion in an English-speaking environment. Both Korean and Spanish speakers have lived in the U.S. for more than three and half years while Mandarin speakers have lived in the U.S. for more than one year.

Table 1 Demographics of the Four Language Groups

	English (N = 35)	Korean (N = 34)	Mandarin (N = 30)	Spanish (N = 23)
Chronological age	20.0 (2.12)	26.5 (5.27)	24.4 (2.69)	27.6 (4.73)

Cloze test (out of 50)	46.3 (2.25)	33.9 (6.92)	37.8 (3.69)	39.3 (5.45)
Self-rated understanding spoken language (out of 10)	9.73 (.449)	6.61 (1.52)	7.07 (1.41)	8.12 (.833)
Age of L2 acquisition	N/A	10.8 (2.73)	9.6 (2.97)	9.0 (5.60)
Length of U.S. residence (years)	N/A	3.63 (2.72)	1.39 (1.37)	3.77 (2.48)
Percentage of L2 use (out of 100)	N/A	45.3 (16.2)	42.9 (12.6)	51.8 (19.1)

Note. Standard deviations in parenthesis.

Native English speakers were significantly younger and had significantly higher cloze test scores and self-rated proficiency than the three L2 groups (all *p*s < .001). One-way ANOVA showed that the three L2 groups did not differ significantly in AOA and amount of L2 use ($F(2, 83) = 2.156$, $p = .122$; $F(2, 85) = 1.678$, $p = .193$, respectively). However, there was a significant difference in their chronological age ($F(2, 83) = 3.791$, $p = .026$), length of U.S. residence ($F(2, 83) = 9.919$, $p < .001$), cloze test scores ($F(2, 83) = 7.522$, $p = .001$), and self-rated proficiency in understanding spoken language ($F(2, 83) = 9.554$, $p < .001$). A series of independent sample *t*-tests were conducted to examine the between-group differences in these four areas. The Mandarin group scored significantly higher in the cloze test ($t(62) = 2.846$, $p = .006$) and were significantly younger than the Korean group ($t(61) = -2.048$, $p = .046$), although the Korean group has lived in the U.S. significantly longer than the Mandarin group ($t(61) = -4.236$, $p < .001$). There was no significant difference in their self-rated proficiency in understanding ($t(60) = 1.237$, $p = .221$). Spanish speakers rated their understanding proficiency significantly higher ($t(52) = -3.383$, $p = .001$), had lived in the U.S. significantly longer ($t(52) = -4.271$, $p < .001$), and were significantly older than Mandarin speakers ($t(52) = -3.013$, $p = .005$), although the two groups did not differ significantly in their cloze test scores ($t(52) = -1.204$, $p = .236$). Finally, Spanish speakers had significantly higher cloze test scores ($t(54) = -$

3.216, $p = .002$) and higher self-rated proficiency in understanding ($t(54) = -4.844$, $p < .001$) than Korean speakers, even though the two groups did not differ significantly in age ($t(55) = -.835$, $p = .407$) or length of U.S. residence ($t(55) = -.202$, $p = .841$). Since English proficiency could not be matched among the three L2 groups and there could be cultural differences in self-rating, scores from the subjective cloze test would be treated as a covariate in all subsequent analyses to account for the variance in L2 proficiency.

Participants in the Korean and Mandarin groups were tested in Los Angeles at the University of Southern California (USC) while participants in the English and Spanish groups were tested at the University of Maryland, College Park. The protocol of the current dissertation has been approved by the University of Maryland Institutional Review Board (UMD IRB Protocol # 11-0679). The IRB at both UMD and USC have an institutional agreement which authorizes data collection under the current protocol at USC.

Chapter 4 - Experiment 1 Stress Cues versus Lexical Knowledge

This experiment compared the use of stress and lexical knowledge by adapting the cross-modal priming lexical decision task from Mattys et al. (2005, Exp.3). However, stimuli were entirely re-selected for the purpose of this study. Participants listened to a five-syllable phrase (e.g., anythingcorri) with visual presentation of a three-syllable letter string (e.g., corridor) 100ms after the offset of the auditory phrase. The first three syllables in the auditory phrase (e.g., *anything*) were referred to as the context and the last two syllables were referred to as the prime (e.g., *corri*). The participants' task is to determine whether the visual target is a real English word. The design was 2 × 2 × 2 factorial with the three factors being condition (primed: *corri* was congruent with target *corridor* or baseline: *corri* was replaced by distorted speech and not congruent with target), lexicality of context (real word or nonword), and stress pattern of prime (strong-weak or weak-strong). This was done separately for each group, rather than language group being a fixed factor. The purpose of using a priming paradigm was to examine the speed of implicit processing. Priming effects were measured as the difference in response time (RT) between the primed and baseline conditions. The use of stress in segmentation was operationalized as a difference in the magnitude of priming effects to one stress pattern compared to another stress pattern in the primes (e.g., SW vs. WS: *regis* vs. *remem*). The use of lexical knowledge in segmentation was operationalized as the difference in the magnitude of priming effects to target words preceded by real word contexts compared to those preceded by nonword contexts (e.g., *considerregis* vs. *dilicterregis* for the target word register).

Materials and Design

Ninety-six pairs of initial-stressed and medial-stressed trisyllabic monomorphemic words were selected from a list generated via the English Lexicon Project (Balota et al., 2007). Each word in the pair was closely matched on surface frequency based on the norms from the Hyperspace Analogue to Language (HAL) database (Lund & Burgess, 1996). This list was sent to 10 Mandarin, 10 Korean, and 10 Spanish L2 learners of English in the same population where the current sample was drawn from. The raters were asked to rate how familiar they were to each word based on a 7-point Likert scale with 1 being “not familiar at all” and 7 being “very familiar”. Only words with a mean familiarity rating higher than 5 were selected to ensure that the L2 learners would not treat them as nonwords. The final stimuli list consisted of 20 pairs of words matched on the rhyme of the final syllable that served as contexts and 20 pairs of words matched on the onset of the first syllable that served as primes. Each context-prime pair was not semantically related. Each word in the pair was matched on written and spoken frequency, familiarity, number of letters, number of phonemes, biphone frequency (both token and type), phonological neighborhood, and uniqueness point (Table 2).

Nonword contexts were created using the Phonotactic Probability calculator (Vitevitch & Luce, 2004) so that each phoneme in the nonword is matched with the phoneme in the corresponding real word in terms of position-specific probability. Twenty pairs of nonwords were created and each pair consisted of an initial-stressed word and a medial-stressed word.

Table 2 Characteristics of the Stimuli in Experiment 1

	Context		Target	
	SW (*character*)	WS (*consider*)	SW (*register)*	WS (*remember)*
Written frequency (log)	8.79	8.93	7.73	7.95
Spoken frequency (log)	2.54	2.59	2.09	2.24
Number of letters	7.7	8.0	7.9	7.85
Number of phonemes	6.75	6.75	7.05	7.3
Familiarity (out of 7)	6.67	6.72	6.65	6.58
Biphone frequency (token)	322.67	294.25	152.8	265.57
Biphone frequency (type)	35.78	31.95	25.18	25.02
Phonological neighborhood size	.45	.15	.3	.25
Uniqueness point	8.7	8.2	8.7	8.3

The design was a 2 x 2 x 2 factorial design with the three factors being condition (whether the auditory prime was congruent with the target or not), lexicailty of the context (real words or nonwords), and stress pattern of the prime (initial-stressed or medial-stressed). There were equal numbers of initial-stressed and medial-stressed words in the context to ensure participants did not develop processing bias for one particular stress pattern. Thus, there were eight possible combinations of the test phrases for each of the 20 pairs of contexts and primes. A test phrase was made by combining the context and the first two syllables of the prime word (e.g., *characterregis*). To prevent repeated exposure to the same prime or context, four test phrases were selected from each set of eight and the selection was counterbalanced on the three factors. Two lists were created so that List 1 contains four test phrases from one set and List 2 contains the other four test phrases from the same set. All participants heard 80 experimental utterances. Appendix A shows a complete list of the critical stimuli from Experiment 1.

The baseline was created by processing the real word primes (e.g., the last two syllable of the test phrase) through a 1-channel noise band vocoder using Tigerspeech (a speech processor that simulates the hearing condition of cochlear-implant users, retrieved from http://www.tigerspeech.com/tst_tigercis.html) to filter out all detailed speech information and maintain the same duration. Two different sets of 80 baseline phrases were created for List 1 and List 2.

One female native American English speaker (with Northeastern dialect) recorded all stimuli. She pronounced each full phrase without interruption (e.g., *considerremember*). Recording was done in a quiet room using an Audio-Technica ATR 20 low impedance microphone. The sounds were recorded using SONY Sound Forge and the files were stored as uncompressed WAV, digitized at 44.1kHz at 16bits. After recording, the last syllable of the phrase was manually cut out, leaving the five-syllable test phrase (e.g., *considerremem*). In each experimental trial, the visual prime (e.g., *remember*) was presented 100ms after the offset of the auditory test phrase (e.g., *considerremem*) or the baseline phrase. To prevent participants from developing processing strategies, three types of fillers were created. The first type of fillers consisted of all nonword visual targets to balance the number of “yes” and “no” responses. Similar to the experimental trials, the nonword targets were presented 100ms after the offset of the auditory phrase. The second type of fillers consisted of half nonword and half real word targets which were presented immediately after the offset of the third syllable in the auditory phrase. The third type of fillers consisted of half real word and half nonword targets which were presented immediately after the offset of the second syllable in the auditory phrase. There were 160 trials in the first type of fillers and 100 trials in each of the second and third types of fillers. All filler trials were equally divided between those with real word or nonword contexts and those with congruent or incongruent primes.

Experiment 1 consisted of a total of 520 trials, equally divided into four blocks of 130 trials each. The presentation of trials in each block was pseudorandomized so that there were at least 40 trials between two repeated primes or contexts. It was ensured that there were no more than three consecutive real word or nonword targets in a row. The order of blocks was counterbalanced across participants.

Procedure

Participants were tested individually using a desktop or laptop PC in a quiet room. They were randomly assigned to List 1 or List 2. Each participant completed all four blocks with a 5-minute break between each block. The experiment was implemented via the E-prime software (Psychology Software Inc., Pittsburgh, PA) which recorded participants' response accuracy and latency. In each trial, participants heard the auditory phrase over high quality SONY headphones and then saw the sequence of letter strings visually presented on the center of the screen written in 22 pt bold Courier font. Participants were instructed to decide whether the letter strings constitute a real English word by pressing the keys labeled "Yes" or "No" on the computer keyboard. Participants were asked to response as accurately and quickly as possible. Participants completed practice trials with feedback before the actual experiment to familiarize with the procedure.

Hypotheses

Accuracy was expected to be high for all language groups since all critical stimuli have high frequency and familiarity so that speakers (especially the L2 learners) do not treat them as nonwords. All language groups were expected to show stronger priming effects for target words preceded by real word than nonword contexts, which would be evidence for the use of lexical knowledge in segmentation. However, the size of the stress effect were expected to be greater than the size of the lexical effect for Mandarin and Spanish speakers whereas the opposite were expected to be found for English speakers (Table 3). Mandarin speakers were expected to be more likely to show a stronger stress effect than Spanish speakers given that Mandarin is more similar to English in terms of the predominance of word-initial stress. Korean speakers were expected to be less likely to use word-level stress as a segmentation cue, thus the magnitude of priming effects were expected to not differ between SW and WS words.

Table 3 Predictions (larger numeric magnitude represents greater weight given to that cue)

	English	Korean	Mandarin	Spanish
Lexical knowledge	2	1	1	1
Stress cue	1	0	2	2

Data Analyses

The data from 30 native English speakers, 34 Korean speakers, 30 Mandarin speakers, and 23 Spanish speakers were analyzed. All analyses were carried out in R studio, an open source programming environment for statistical computing (R Development Core Team, 2007) with the lme4 package for linear mixed-effectss modeling (Bates, 2005). Analyses were based on unaveraged and untrimmed data. Response time (RT) data for incorrect responses were excluded (4.35% of the total data). All RT data were log-transformed to improve normality. For the RT model *p*-values were based on MCMC sampling (Baayen, 2008; Bates, 2005) and for the accuracy model, they are based on the *Wald z* distribution. The same analytical method was applied in all subsequent analyses in Experiments 2-4.

In the current experiment, to examine the effects of lexical knowledge and stress location on priming effects (e.g., RT differences between baseline and primed trials), reliable interactions between priming conditions and the two factors would be the main focus of the analyses. A statistically significant interaction between priming conditions and stress patterns would be evidence for the use of stress cues in segmentation because this result indicates that different stress location (initial vs. medial) affects the magnitude of priming effect and larger priming effects indicate faster segmentation. Data were analyzed in a linear mixed-effects model with priming conditions, stress patterns, context lexicality, and cloze test scores (as a covariate for proficiency) as fixed effects and subject and item as random effects for each individual language group. Accuracy rates were analyzed using a similar model but with the binomial function. Previous studies examining native English segmentation (e.g. Mattys et al. 2005; Cutler & Norris, 1988) did not include English proficiency in their analyses. In the current dissertation, cloze test scores were included as a covariate in data analysis of the English group to keep the model consistent across all four language groups. Another set of mixed-effects models with language group as a fixed effect is shown in Appendix E. Block and list were entered as fixed effects in the initial model but were removed in the final model since they did not contribute to the best goodness-of-fit in model comparison.

Results

Table 4 Mean lexical decision RTs and % accuracy (in parentheses). Priming effect is the difference in RT between the baseline and primed conditions.

	English		Korean		Mandarin		Spanish	
Prime	SW	WS	SW	WS	SW	WS	SW	WS
Word context	543 (97.1)	528 (97.8)	875 (95.3)	859 (95.3)	836 (98.3)	829 (91.2)	668 (96.3)	671 (96.9)
Baseline	573 (94.8)	580 (96.0)	886 (94.7)	876 (94.3)	836 (95.5)	871 (95.5)	698 (97.7)	722 (97.7)
Mean priming	30	52	11	17	0	42	30	51
Nonword context	558 (97.1)	534 (97.0)	879 (95.5)	858 (94.3)	840 (95.3)	808 (92.7)	673 (95.4)	686 (97.5)
Baseline	577 (97.3)	571 (97.3)	864 (93.9)	931 (92.2)	870 (93.5)	843 (94.3)	687 (96.7)	684 (97.7)
Mean priming	19	37	-15	73	30	35	14	-2

Table 4 shows the mean lexical decision latencies and accuracy rates. For brevity, only interactions between priming condition and segmentation cues (e.g., context lexicality and stress pattern) are reported since reliable interaction reflects the effect of lexical knowledge or stress location on priming. The main findings from this experiment (Figure 3) were that when the target words were preceded by real word context words, English and Spanish speakers showed greater priming effects for medial-stressed (WS) primes than for initial-stressed primes (SW). In contrast, Korean speakers showed this difference in the priming effects of SW and WS primes with nonword contexts but not with real word contexts. Korean and Spanish speakers showed a trend of cue interaction whereas native English speakers appeared to use the two types of cues independently. Mandarin speakers did not show any significant effect or interaction.

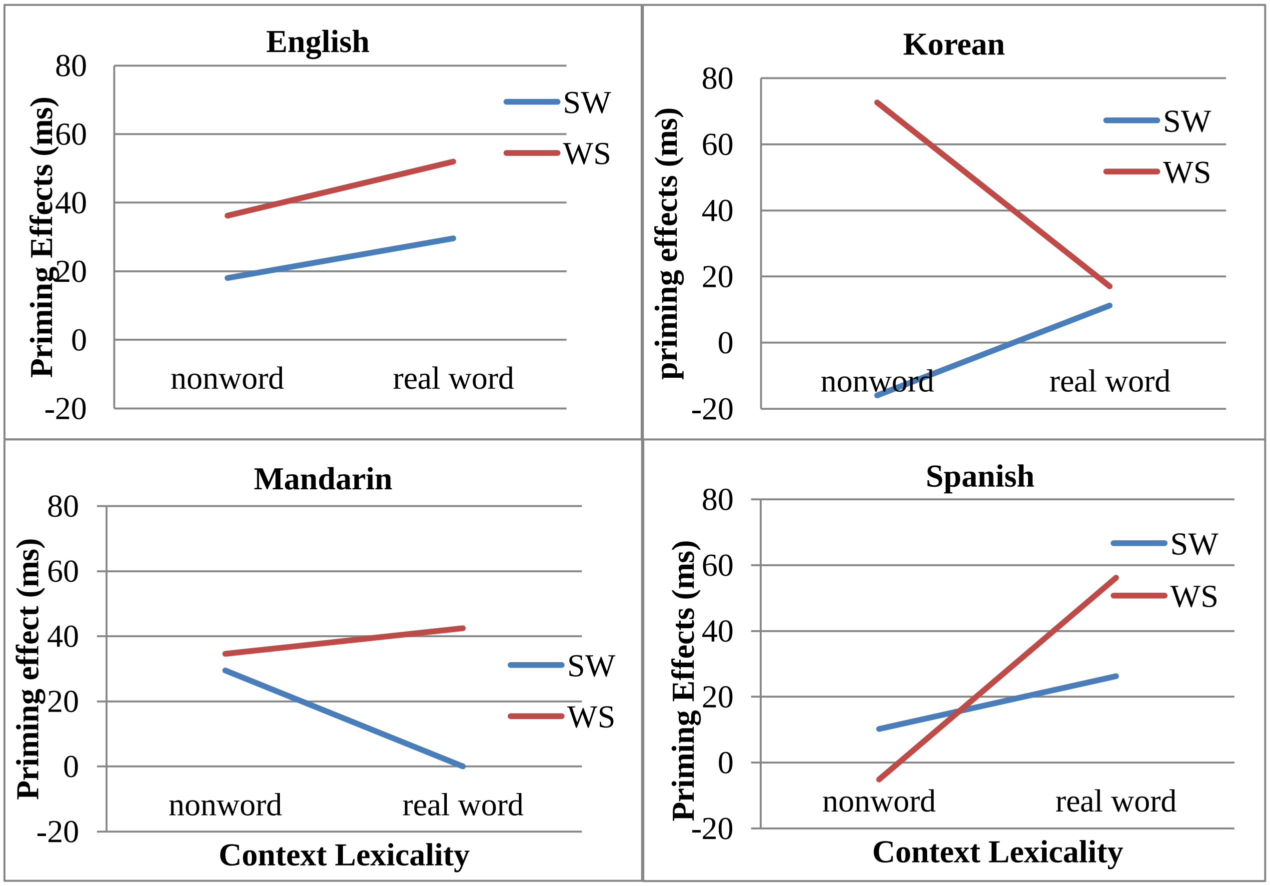

Figure 3 The use of stress cues and lexical knowledge (priming effects in RT)

English

Response Times. In the model predicting response latency (Table 5), the intercept estimated English speakers' log RT for target words in the baseline condition, with nonword contexts and initial-stressed primes. There was a significant interaction between priming condition and context lexicality. English speakers showed greater priming effects when the target word was preceded by real word context than when it was preceded by nonword context for initial-stressed primes (Figure 3). There was also a significant interaction between priming condition and stress location. Priming effects were significantly larger when prime was medial-stressed than when it was initial-stressed for nonword context. The three way interaction among priming condition, lexicality, and stress was not significant. There was a significant effect of cloze test scores, showing higher English proficiency was associated with faster RT in the lexical decision task (Figure 4). This result suggests that even for native speakers, English proficiency influences the speed of word recognition.

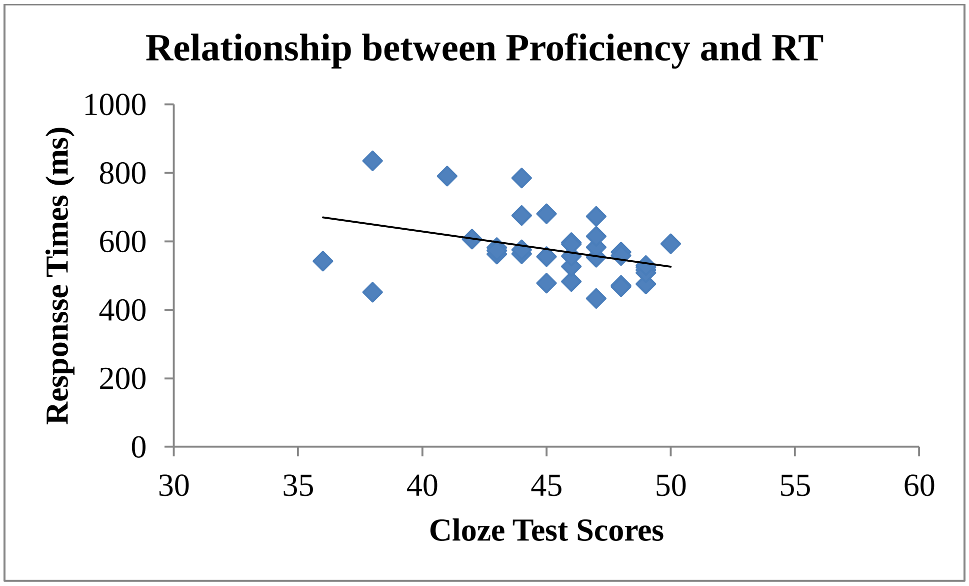

Figure 4 The significant association between proficiency and RT in the English group

Accuracy. In the model predicting response accuracy, none of the interactions involving priming condition were statistically significant (all *p*s >.1).

Table 5 Mixed-effects linear model estimating log RT in the English group.

Fixed Effects	Estimate	Std. Error	*t*-value	*p*MCMC
Intercept	3.039	.141	21.575	.0001
Condition (primed)	-.012	.008	-1.57	.123
Context (real word)	.003	.005	.532	.583
Stress (WS)	.0004	.005	.081	.934
Cloze	-.006	.003	-2.125	.014
Priming × Lexicality	-.016	.007	-2.206	.029
Priming × Stress	-.016	.007	-2.219	.025
Priming × Lexicality × Stress	-.001	.01	-.119	.896

Korean

Response Times. There was a significant interaction between priming condition and stress location (Table 6). Korean speakers showed greater priming effects for WS words than for SW words when primes were preceded by nonword context. More importantly, there was a significant three way interaction among priming conditions, context lexicality, and stress pattern, showing that the use of stress cues is influenced by context lexicality. When context was a real word, the difference in the magnitude of priming effects between SW and WS primes are small; when context was a nonword, priming effect was larger for WS primes compared to SW primes (Figure 3). There was a significant effect of cloze test scores; higher English proficiency was associated with faster RT.

Accuracy. In the model predicting response accuracy, none of the interactions involving priming condition were statistically significant (all *p*s > .1). There was a significant effect of cloze test scores ($z = 3.984$, $p < .001$), suggesting that higher L2 proficiency is associated with higher accuracy in making lexical judgments.

Table 6 Mixed-effects linear model estimating log RT in the Korean group.

Fixed Effects	Estimate	Std. Error	*t*-value	*p*MCMC
Intercept	3.073	.087	35.37	.0001
Condition (primed)	.003	.013	.20	.857
Context (real word)	.011	.008	1.54	.130
Stress (WS)	.030	.008	3.97	.0001
Cloze	-.005	.002	-2.04	.017
Priming × Lexicality	-.008	.011	-.76	.452
Priming × Stress	-.039	.011	-3.69	.0004
Priming × Lexicality × Stress	.029	.015	1.97	.043

Mandarin

Response Times. Neither context lexicality nor stress location interacts with priming condition significantly in the Mandarin group (Table 7). The only reliable finding was the significant effect of cloze test scores which suggests higher English proficiency was associated with faster RT in making lexical judgments.

Table 7 Mixed-effects linear model estimating log RT in the Mandarin group.

Fixed Effects	Estimate	Std. Error	*t*-value	*p*MCMC
Intercept	3.183	.122	25.989	.0001
Condition (primed)	-.017	.013	-1.333	.189
Context (real word)	-.008	.008	-1.028	.308
Stress (WS)	-.008	.008	1.083	.273
Cloze	-.007	.003	-2.259	.018
Priming × Lexicality	.008	.011	.700	.489
Priming × Stress	-.002	.011	-.167	.868
Priming × Lexicality × Stress	-.011	.016	-.710	.476

Accuracy. In the model predicting response accuracy, none of the interactions involving priming condition reached statistical significance (all *p*s > .1)

Spanish

Response Times. There was a marginally significant three-way interaction among priming condition, context lexicality, and stress location (Table 8). As Figure 3 shows, when context was a nonword, the difference in the magnitude of priming effects was small between initial-stressed and medial-stressed primes; when context was a real word, priming effects was greater for medial-stressed primes than for initial-stressed primes.

Accuracy. In the model predicting response accuracy, none of the interactions involving priming condition were statistically significant (all *p*s > .1).

Table 8 Mixed-effects linear model estimating log RT in the Spanish group.

Fixed Effects	Estimate	Std. Error	*t*-value	*p*MCMC
Intercept	2.972	.127	23.340	.0001
Condition (primed)	-.008	.010	-.901	.367
Context (real word)	.002	.007	.243	.797
Stress (WS)	-.002	.007	-.234	.833
Cloze	-.004	.003	-1.274	.118
Priming × Lexicality	-.003	.010	-.254	.798
Priming × Stress	.004	.010	.391	.687
Priming × Lexicality × Stress	-.028	.014	-1.929	.052

Discussion

In summary, English speakers showed faster segmentation when the context word was a real word than when it was a nonword. This result provided evidence for the use of lexical knowledge in native English segmentation and this is consistent with the hypothesis. Also, they were faster to segment medial-stressed primes than initial-stressed primes and this result was not consistent with the hypothesis. The absence of a significant interaction between these stress location and context lexicality suggests that English speakers use lexical knowledge independent of stress location. Both Korean and Spanish speakers showed a trend of cue interaction in L2 segmentation, albeit in opposite directions. Korean speakers were faster to segment medial-stressed primes than initial-stressed primes when the context was a nonword; this difference in segmentation latency between the two stress patterns was significantly smaller when the context was a real word. In contrast, Spanish speakers were faster to segment medial-stressed primes than initial-stressed primes when the context was a real word; the difference in segmentation latency between the two stress patterns was smaller when the context was a nonword. Mandarin speakers did not show any significant effect or interaction involving priming conditions.

Stress Cues

English, Korean, and Spanish speakers showed faster segmentation of medial-stressed words compared to initial-stressed words. However, the stress effect did not interact with the context lexicality effect in the English group whereas it did in the Spanish and Korean groups. The fact that the native and nonnative groups showed different patterns of results suggest that there may be influence from L1 prosodic transfer. For Spanish speakers, the WS stress effect may be interpreted as cross-linguistic influences on the use of stress location in L2 segmentation. Stress mostly falls on the penultimate syllable in Spanish (Navarro, 1966). Medial-stress in trisyllabic words, which were the stimuli used in the current experiment, would be consistent with the predominant stress pattern in Spanish speakers' L1. Thus, Spanish speakers' use of stress cues in L2 segmentation may be influenced by L1 stress characteristics.

Korean speakers showed a significant WS stress effect in nonword context. The use of stress cues by Korean speaker is not consistent with the hypothesis. Since lexically contrastive stress does not exist in Korean and previous research has shown that Korean speakers do not encode stress in lexical access (Lin et al., under review), it was hypothesized that Korean speakers would not be able to use stress location as a cue to word boundary in L2 segmentation. However, recent research has shown that speakers whose L1 does not have lexically contrastive stress were able to use prominence as a cue in speech segmentation (Kim, Broersma, & Cho, 2012; Tremblay, Coughlin, Bahler, & Gaillard, 2012).

In languages without contrastive stress such as Korean and French, not every syllable is given equal prominence and speakers of these languages may be sensitive to these prominence distinctions. For example, the Accentual Phrase in standard Korean is marked by a phrase-final High tone (Jun, 1998), which is realized through a rise in fundamental frequency (F0). Research has shown that native Korean speakers could use this change in pitch to identify word boundaries (Kim et al., 2012). In an artificial language (AL) learning task, Korean listeners were found to rely on the F0 rise cue and lengthening cue in word-final position to segment the newly acquired AL sequences. Similarly, native French speakers also use the prominence distinction in speech segmentation (Christophe, Peperkamp, Pallier, Block, & Mehler, 2004; Tremblay, et al., 2012; Tyler & Cutler, 2009). Korean speakers in the current study may be sensitive to the pitch cues in stressed syllables and thus were able to use phonetic prominence in L2 segmentation.

For English speakers, the direction of the priming effect is not consistent with the hypothesis. English speakers were expected to show stronger priming effects for initial-stressed words. Previous research has shown that English speakers were faster to segment words with initial-stressed (Cutler & Norris, 1988; Mattys et al., 2005, Exp3). The segmentation advantage of medial-stressed words is surprising since stress location does not coincide with word edges. One possible explanation is that the stress location in the context words may have influenced English speakers' use of stress cues to segment the target words. Participants may have initially considered the five-syllable auditory phrase such as *considerregis* or *manisterregis* as a whole word. The general stress patterns of pentasyllabic words in English are typically [S W W S W] as in *characteristic* or *qualification* and [W S W S W] as in *enthusiastic* or *consideration,* for both patterns the first S is secondary and the second S is primary. Since the stress pattern of the auditory phrase containing SW primes (i.e. SWWSW or WSWSW) are consistent with the typical stress pattern of pentasyllabic words in English, listeners might be slower to recognize that the five-syllable sequence is consisted of two words, thus delaying the identification of the word boundary and subsequent recognition of the target word. In contrast, the auditory phrase containing WS primes (i.e. SWWWS or WSWWS) are not consistent with the typical stress pattern of pentasyllabic words in English. Thus the entire sequence may be reasonably perceived as one English word and must be broken up. As a result, native English speakers were faster to identify the word boundary and recognize the target words for medial-stressed primes.

To test this hypothesis, stress location in context words were added as a fixed effect in generalized linear mixed-effects models for each language group. Initial models with priming condition, context stress location, prime stress location, context lexicality and proficiency did not showed any significant interaction or effect involving context lexicality, thus this term was removed. A significant three-way interaction among priming condition and stress location of context and prime would indicate the location of stress in the context words indeed influence participants' use of stress cues in segmenting the target words. The three-way interaction was significant for the English group (t = 2.803, p = .006). As Figure 5 shows, when context stress pattern was initial-stress, English speakers showed stronger priming effects for medial-stressed target words compared to initial-stressed target words; when context stress pattern was medial-stress, there was no difference in priming effects between initial-stressed and medal-stressed target words.

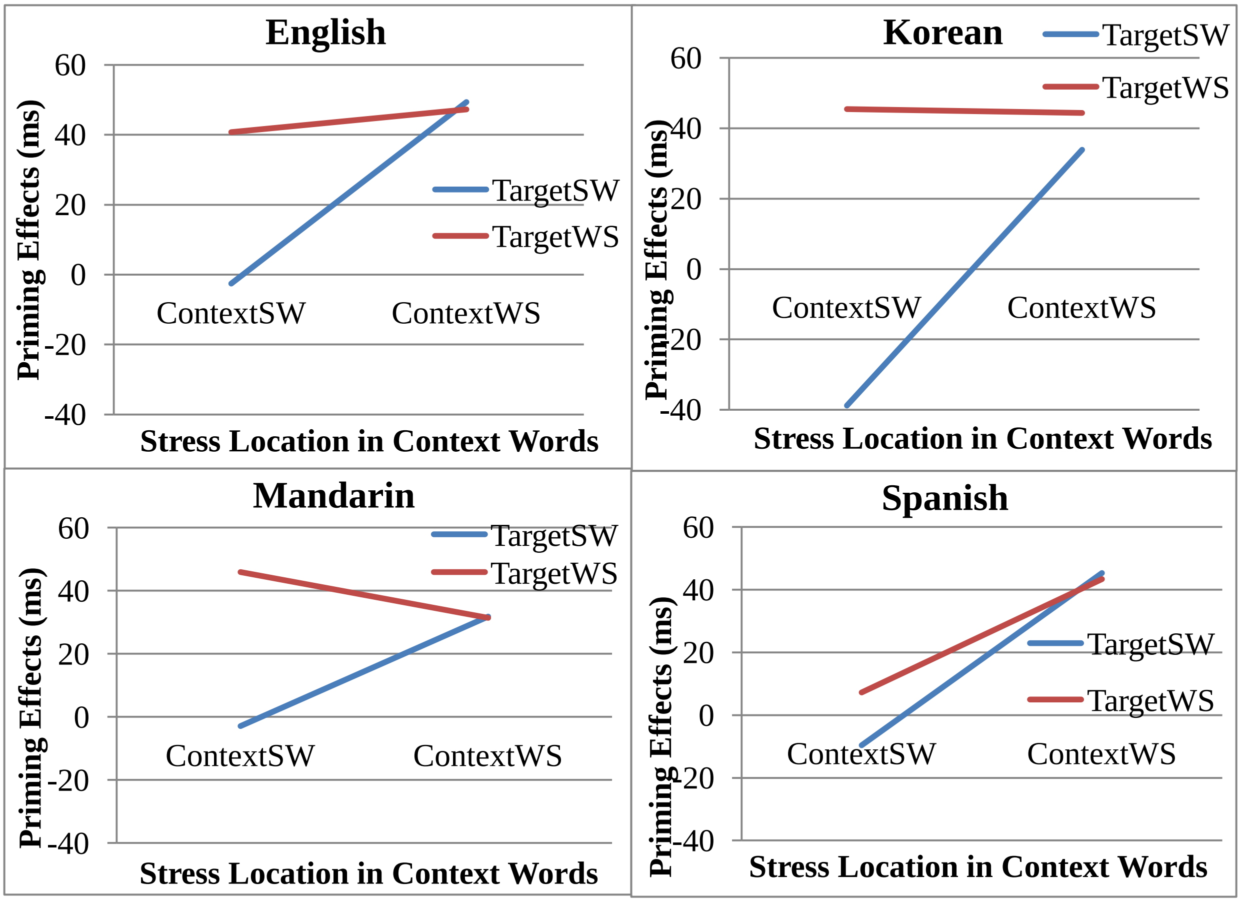

Figure 5 Interaction between context stress location and prime stress location (priming effects in RT)

The five-syllable sequence with the stress pattern SWWWS elicited faster segmentation than the sequence with the pattern SWWSW probably because the former is not a possible stress pattern for pentasyllabic words in English and this encouraged listeners break down the sequence by identifying a word boundary. However, this result could not explain for the finding that WSWSW and WSWWS patterns have similar priming effects. If listeners were indeed trying to parse the word a five-syllable whole word, then WSWWS, which is a typical stress pattern for English pentasyllabic words, should have stronger priming effects than WSWSW. Thus, this hypothesis regarding whole-word parsing cannot adequately explain the WS stress effect.

The post-hoc analysis showed that this three-way interaction among priming condition, context stress pattern, and target stress pattern was not significant for Spanish speakers ($t = .967$, $p = .334$). As Figure 5 shows, medial-stressed words had stronger priming effects than initial-stressed words regardless of the stress pattern of the context word. This finding suggests that Spanish speakers were faster to segment trisyllabic words with medial-stress since this stress pattern is consistent with the predominant stress pattern in their L1, demonstrating the influence of L1 typology on L2 segmentation.

Korean speakers also showed a significant three-way interaction among priming condition, context stress location, and prime stress location ($t = 2.33$, $p = .019$). Korean speakers were faster to segment medial-stressed target words than initial-stressed target words, although this difference was smaller when the context word was medial-stressed. These results suggest that Korean speakers were more sensitive to medial-stressed words overall. Previous research has shown that Korean speakers were able to utilize the rise in pitch to segment artificial language. Since in English stressed syllables have higher pitch than unstressed syllables (Fry, 1955), Korean speakers may be able to use this pitch cue to segment L2 speech even though they do not encode stress in lexical representation (Lin et al., under review). A speculation for the segmentation advantage of medial-stressed words is that having a stressed syllable sandwiched between two unstressed syllables (i.e. WSW) makes the pitch cues more salient because Korean speakers can form two pairs of comparison in relative F0, one between WS and another between SW. In contrast, initial-stressed words (i.e. SWW) only have one point of comparison. Thus, if Korean speakers were relying solely on pitch prominence in stressed syllable to identify word boundaries, the pith cue medial-stressed words may be easier do detect. Given that Korean speakers only showed a significant stress effect in nonword context (Figure 4) whereas the significant stress effect was observed in the English group regardless of context lexicality, it is reasonable to conclude that the Korean speakers' use of stress cues was influenced by L1 prosodic structures.

Lexical Knowledge

English speakers showed stronger priming effects for real word context than nonword context and this lexicality effect was present regardless of the stress location of the primes. This finding suggests that English speakers use lexical knowledge in native segmentation and this is consistent with the hypothesis. Mandarin speakers did not show a significant lexicality effect while Korean and Spanish speakers showed a trend of cue interaction (see the following section). In the English group, the similar significance value of the stress effect (p = .025) and the lexical effect (p = .029) suggests that native English speakers in the current study did not weigh one cue over the other. This is not consistent with the Hierarchical Framework (Mattys et al., 2005) which postulates that under optimal listening condition, English speakers would only use lexical knowledge but not stress location in L1 segmentation. The current results suggest that when the speech signal is clear, English speakers use all of the cues available to them. This is consistent with previous research which has shown that humans are active and opportunistic learners who capitalize on all functional regularities present in the environment (Gomez, 2002; Seligman, 1970). Since the stress and lexical cues are not providing conflicting information about word boundary, it would be beneficial to take advantage of them both to accelerate the segmentation process.

Cue Interaction

Both Korean and Spanish groups showed a trend of cue interaction. Korean speakers were faster to segment medial-stressed words than initial-stressed words when the context was a nonword; yet when the context was a real word, the difference in priming effects between the two stress patterns was significantly smaller. Spanish speakers showed stronger priming effects for medial-stressed words than initial-stressed words when the context was a real word and the difference in priming effect was smaller when the context was a nonword (though the interaction was only marginally significant).

Spanish has lexically contrastive stress and previous research has shown that Spanish speakers encode stress in lexical access (Dupoux et al., 2001). When lexical cues were available (i.e. real word context), Spanish speakers used stress cues in a native-like manner in L2 segmentation; when lexical cues were not available in speech signal (i.e. nonword context), Spanish speakers' use of stress cues did not resemble that of English speakers. On the other hand, it has been suggested that Korean speaker do not encode stress in their phonological representation of L2 English words (Lin et al., under review). It is possible that when Korean speakers encountered real English words, they were not able to use stress cues to identify word boundaries since the abstract representation of stress was not available in lexical access (Lin et al., under review). Participants were asked to listen to the auditory phrase passively and focus their attention on the visual target words which they would make lexical judgments on. Due to the nature of this task, the absence of stress representation may be less critical in the processing of nonword context since lexical access is less likely to be involved. As a result, Korean speakers might have used phonetic prominence (more specifically, pitch cues) in stressed syllables, as opposed to the stress pattern, to identify word boundaries. Korean speakers may weight lexical cues over stress cues since the stress effect only emerged in nonword context when they could not utilize lexical knowledge. In contrast, Spanish speakers show a significant effect when lexical cues are available, suggesting that they do not weight lexical knowledge over stress cues. However, the fact that the L2 groups showed cue integration but the native group did not implies that native and nonnative segmentation involves fundamentally different use of cues.

Chapter 5 - Experiment 2 Phonotactic Cues versus Lexical Knowledge

This experiment compared the use of cues at the segmental level and those at the lexical level. Although belonging to the same segmental level, phonotactic cues and acoustic-phonetic cues entail different levels of knowledge and representation. As hypothesized in RFL2 (Figure 2), the strength of phonotactic constraints on segmentation is weaker than that of junctural or allophonic cues and this prediction is based on findings from native English speakers (Newman et al., 2011). The use of phonotactic cues was operationalized as faster response times (RT) to segment target words preceded by a coda consistent with L1 phonotactic constraint and slower RT for those preceded by a coda that violates L1 phonotactic rules. Similar to Experiment 1, lexical knowledge was operationalized as the recognition time for a target word preceded by a real-word or a nonwod context.

In particular, items with three potential segments at the word boundary, /n, ŋ, s/, were compared. /n/ is a legal onset and coda in all four languages (Table 9), although velarization of final /n/ is frequent in Caribbean and Central American dialects of Spanish (Kochetov & Colantoni, 2011). /ŋ/ can only occur syllable-finally in English, Korean and Mandarin. Since the probability of a word boundary following /ŋ/ is relatively higher than that following /n/, English, Korean, and Mandarin speakers may be faster to segment target words preceded by /ŋ/ than those preceded by /n/. Spanish speakers may be slower to segment target words preceded by /ŋ/ than those preceded by /n/ since /ŋ/ does not exist in the Spanish phoneme inventory. Mandarin and Korean do not allow /s/ in the syllable-final position. Mandarin and Korean speakers may take longer to recognize target words preceded by/s/ since they are more likely to erroneously segment /s/ as the onset of the following word at first parsing. In contrast, /s/ is legal coda in English and English speakers' segmentation latency should not differ for target words preceded by /n/ or /s/ if English speakers are only sensitive to phonotactic legality. Alternatively, if English speakers are sensitive to phonotactic probability, then target words preceded by /n/ should be identified faster than those preceded by /s/ given that /s/ has a higher frequency of occurrence in word-final than in word-initial positions. Although /s/ is allowed syllable-finally in Spanish, word-final /s/ has more phonological variations than word-final /n/ and Boomershine (2006) has shown that native Spanish speakers have more difficulty recognizing L1 words with /s/ coda than those with /n/ coda. If this difficulty with word-final /s/ has cross-linguistic variations on the segmentation of L2 words, Spanish speakers may take longer to segment target words preceded by coda /s/.

Unlike Mandarin which only allows two phonemes in syllable coda, Korean has a slightly larger coda inventory, allowing /p, t, k, m, l/ at syllable-final positions in addition to /n, ŋ/. Thus, the probability of a word ending with /n/ in Mandarin is higher than that in Korean. Mandarin speakers may rely more on coda cues than their Korean counterparts. Unfortunately, these fine-grained cross-linguistic variations are beyond the scope of this project.

Table 9 Typological differences in phonotactics (absolute legality)

	English		Mandarin		Korean		Spanish	
	onset	coda	onset	coda	onset	coda	onset	coda
/n/	√	√	√	√	√	√	√	√
/ŋ/	×	√	×	√	×	√	×	×
/s/	√	√	√	×	√	×	√	√

Materials and Design

This experiment employed a word spotting task, commonly used in previous literature (e.g., Endress & Hauser, 2010; Mattys et al., 2005; Vroomen et al., 1998). In each trial, a visual target appears on the screen for 1000ms, following by the immediate presentation of an auditory phrase. Participants were asked to decide whether the auditory phrase contains the target word they have seen before. The auditory phrase consists of six syllables in which the first three syllables constitute the *auditory context* and the last three syllables make up the *auditory target*. The design of this experiment is 2 × 3 factorial with the two factors being lexicality of the contexts (real words or nonwords) and syllable codas of the contexts (/n ŋ s/).

An initial list of 90 trisyllabic words was generated via the English Lexicon Project (Balota et al., 2007) based on their matched written frequency from the HAL database (Lund & Burgess, 1996). Thirty of the 90 words have /n/ as a coda, thirty of them ends with /ŋ/ and thirty of them with /s/. Familiarity ratings were collected from the same group of raters from Experiment 1. The rating was done with a 7-point Likert scale with 1 being "not familiar at all" and 7 being "very familiar". Only words with a mean familiarity rating higher than 6.5 were selected. For words in the coda /s/ condition, it was also ensured that the /s/ is not realized as /z/ word-finally. The final stimuli list consisted of 10 pairs of words, half of them initial-stressed and half medial-stressed, in each of the three coda conditions. Each pair of words was matched on written and spoken frequency, familiarity, number of letters, number of phonemes, the size of phonological neighborhood and uniqueness point (Table 10). However, the /ŋ/ condition has significantly higher biphone token frequency and biphone type frequency than the /s/ and

/n/ conditions ($F(2, 57) = 23.881$, $p < .001$; $F(2, 57) = 56.446$, $p < .001$, respectively), probably due to the highly frequent co-occurrence of the biphone /ɪŋ/ as the English present tense marker *–ing*. Nonwords designed to match the real word contexts were created using the same method as that in Experiment 1 using the Phonotactic Probability calculator (Vitevitch & Luce, 2004). Thirty pairs of nonwords were created, 10 pair in each of the coda condition. The nonwords had the same stress pattern as their corresponding real words. Visual targets were 20 vowel-initial words to make segmentation more difficult and to ensure the word-final /n/ is not velarized and realized as /ŋ/ by the Spanish speakers. These target words were selected from an original list of 30 words and 10 words with a familiarity rating lower than 6.5 were excluded.

Table 10 Mean Statistics of the stimuli

	Coda		
	n	ŋ	s
Written frequency (log)	9.55	9.48	9.52
Spoken frequency (log)	2.67	2.63	2.58
Number of letters	8.5	8.85	8.9
Number of phonemes	7.2	7.2	7.65
Familiarity (out of 7)	6.84	6.87	6.76
Biphone frequency (token)	247.12	819.15	241.57
Biphone frequency (type)	23.84	119.56	27.51
Phonological neighborhood size	.4	.5	.4
Uniqueness point	10.4	10.6	9.95

As a result of the 2 × 3 factorial design, each target was paired with six different contexts for the auditory phrase. To prevent repeated exposure to the same target, three phrases were selected from each set of six and the selection was counterbalanced based on a Latin-square design. Two lists were created so that List 1 contained three phrases from one set and List 2 contained the other three phrases from the same set. There were a total of 60 critical trials in each list. Appendix B shows a complete list of the stimuli from Experiment 2.

To prevent participants from developing processing strategies for the specific codas or only focusing on the last three syllables of the auditory phrase, three types of fillers were created, with 60 trials in each type. For the first type of filler trials, the visual target matched neither the auditory context nor the auditory target. Half of the auditory contexts were real words and half of them were nonwords while the auditory targets were all real words. For the second type of filler trials, the visual target matched the context in the auditory phrases. All auditory contexts were real words while half of the auditory targets were nonwords and half of them were real words. Finally, for the third type of filler trials, the visual target matched neither the auditory context nor the auditory target. All auditory contexts were real words while half of the auditory targets were nonwords and half of them were real words. There were a total of 240 trials, 60 critical trials and 180 filler trials, with an equal number of positive and negative responses. The trials were pseudo-randomized so that there were at least 70 trials separating the same context. It was also ensured that there was no more than three “Yes” or “No” responses consecutively. The sounds were recorded by the same female native English speaker in the same manner as in Experiment 1.

Procedure

Participants were tested individually using a desktop or laptop PC in a quiet room. They were randomly assigned to List 1 or 2. The experiment was implemented via the E-prime software (Psychology Software Inc., Pittsburgh, PA). For each trial, participants first saw a fixation “+” in the center of the screen for 500ms. Then they saw the target word which stays on the screen for 1000ms. Immediately following the visual word, participants heard the auditory phrase and they were instructed to decide whether the auditory phrase contains the visual target word by pressing the keys labeled “Yes” or “No” on the computer keyboard. Speed and accuracy were emphasized. The inter-trial interval was 1000ms. Participants completed eight practice trials with feedback before the actual experiment to familiarize with the procedure.

Hypotheses

Target words preceded by /ŋ/ were expected to result in faster RT, compared to those preceded by /n/, for the English, Mandarin, and Korean groups (Table 11). However, target words preceded by /ŋ/ were expected to result in longer RT for the Spanish speakers. Words preceded by /s/ were hypothesized to result in longer RT for Mandarin and Korean speakers. If phonological variation of word-final /s/ in L1 Spanish influences L2 segmentation, Spanish speakers were hypothesized to be slower to identify target words preceded by /s/. If phonotactic probability influences native segmentation, English speakers were hypothesized to take longer to segment words with coda /s/ since /s/ has similar frequency of occurrence in word-initial and word-final positions in English whereas /n/ has a higher likelihood of occurrence in word-final position. However, if only absolute legality matters in segmentation, both English and Spanish speakers were hypothesized to show similar RT to words with coda /s/ and /n/. All language groups were expected to respond faster to target words preceded by real word contexts than those preceded by nonword contexts. The native speakers were hypothesized to show a stronger lexical effect whereas the L2 learners will show a stronger phonotactic effect.

Table 11 Example stimuli for context and predictions

	coda /n/		coda /ŋ/		coda /s/	
	real word	Nonword	real word	nonword	real word	nonword
	everyone	*Akluben*	*following*	*telewing*	*Evidence*	*delbiens*
English	Baseline	Baseline	Shorter	Shorter	Same/longer	Same/longer
Mandarin	Baseline	Baseline	Shorter	Shorter	Longer	Longer
Korean	Baseline	Baseline	Shorter	Shorter	Longer	Longer
Spanish	Baseline	Baseline	Longer	Longer	Same/longer	Same/longer

Data Analyses

Since three Korean participants and one Spanish participant did not return for session 2, data from 31 Korean, 22 Spanish, 30 Mandarin, and 35 English speakers were analyzed. The dependent variables were accuracy and (log-transformed) RT in the word spotting task. RT data for incorrect responses were excluded (4.63% of total data). The independent variables were the lexicality and codas of the auditory context. Since coda /n/ serves as the baseline of comparison, it was treated as the intercept in the linear mixed-effects model for each language group. Coda, context lexicality, and cloze test scores (as a covariate for proficiency) were fixed effects while subject and item were random effects. Multiple comparisons of means were conducted using the Simultaneous Tests for General Linear Hypotheses from the multcomp package (Hothorn, Bretz, & Westfall, 2008) with Tukey contrasts and adjusted *p*-values. Accuracy rates were analyzed using a similar model but with the binomial function. Results from another set of mixed-effects linear model with language group as an additional fixed effect are shown on Appendix F.

Results

Table 12 shows the mean lexical decision latencies and accuracy rates. The main findings of this experiment (Figures 6 and 7) were that English, Mandarin, and Korean speakers showed an effect of coda /s/ in which they were slower and less accurate to identify target words preceded by coda /s/ compared to those preceded by /n/. This effect was not observed in the Spanish group. In addition, English speakers showed evidence of using lexical knowledge in the /ŋ/ condition.

Table 12 Mean word spotting RTs and % accuracy (in parentheses)

	English		Korean		Mandarin		Spanish	
Context	Word	Non-word	Word	Non-word	Word	Non-word	Word	Non-word
Coda /n/	1121 (96)	1158 (94)	1284 (97)	1301 (95)	1243 (99)	1259 (96)	1266 (97)	1304 (97)
Coda /ŋ/	1107 (98)	1189 (95)	1257 (97)	1299 (97)	1213 (97)	1279 (95)	1252 (95)	1310 (97)
Coda /s/	1171 (97)	1244 (90)	1320 (91)	1389 (92)	1291 (94)	1344 (91)	1315 (98)	1358 (94)

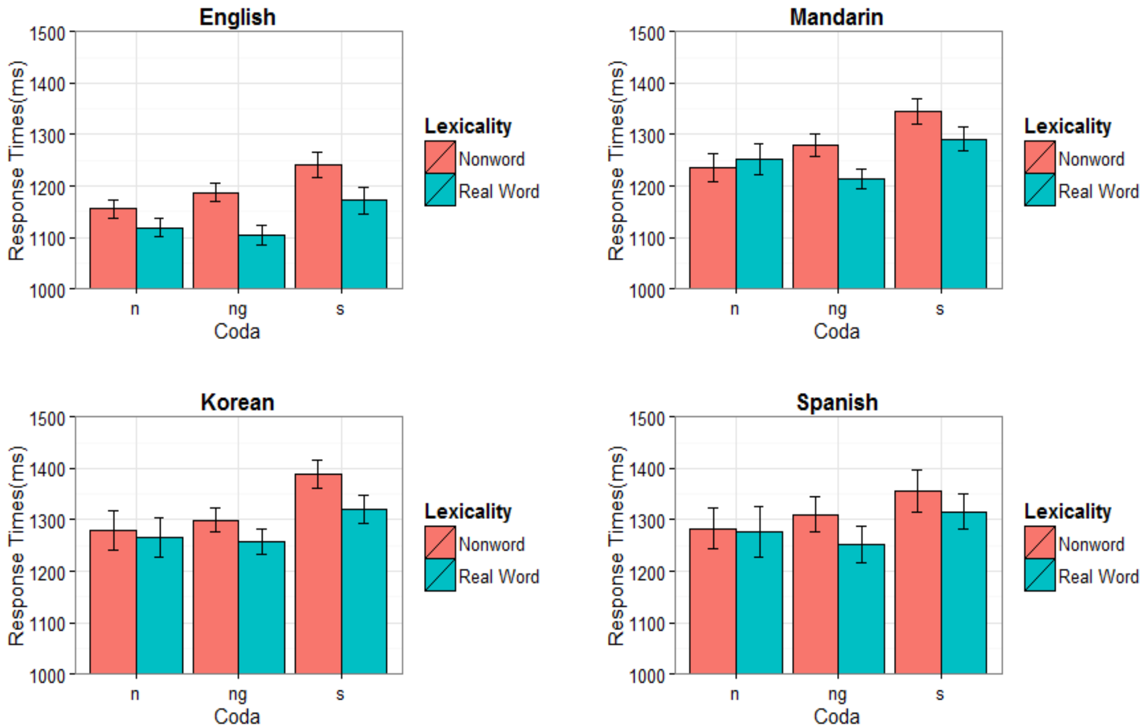

Figure 6 Phonotactic cues vs. lexical knowledge (RT).

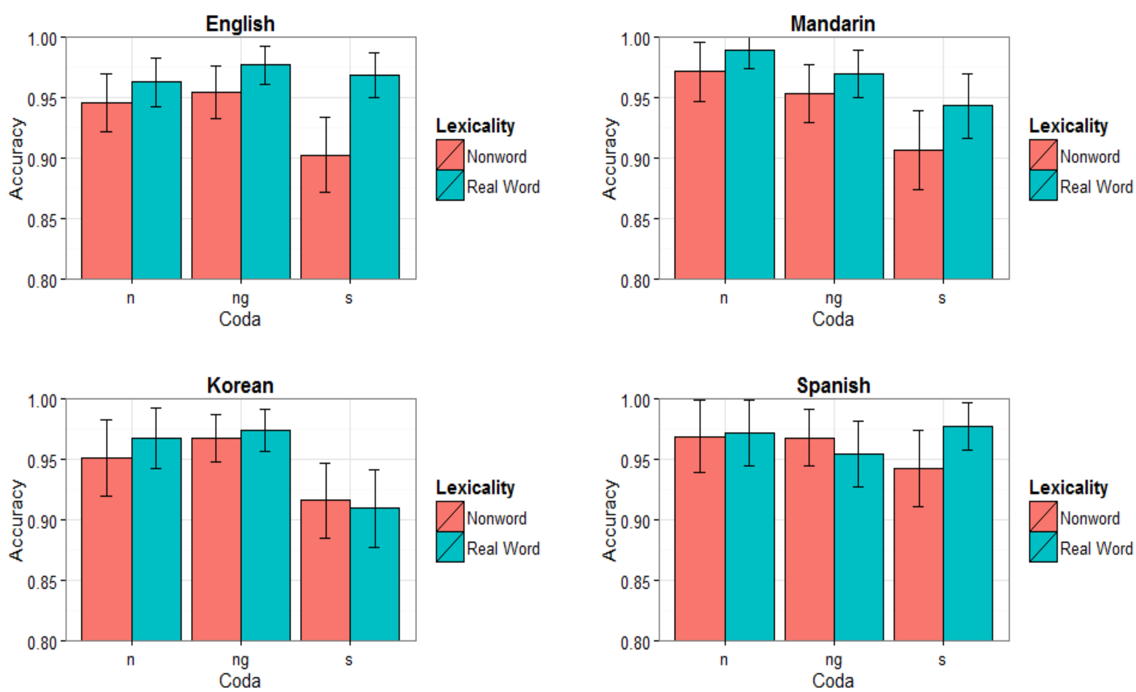

Figure 7 Phonotactic cues vs. lexical knowledge (accuracy).

English

Response Times. In the model predicting response latency for the English group, the intercept estimated English speakers' log RT when the target words were preceded by a nonword context with coda /n/. The model only showed a significant effect of coda /s/ (Table 13). Native speakers were significantly faster to identify target words preceded by coda /n/ compared to those preceded by coda /s/ in nonword contexts (Figure 5). This effect was not significant in real word contexts (z = 1.932, p = .382) based on pairwise comparison. Pairwise comparisons showed a significant lexicality effect in the /ŋ/ condition (z = -3.46, p = .007); English speakers were faster to identify target words preceded by real word contexts than those preceded by nonword contexts. Pairwise comparison did not show any additional significant effect (all *p*s > .1).

Accuracy. In the model predicting accuracy, neither effect nor interaction was significant.

Table 13 Mixed-effects linear model estimating log RT in the English group

Fixed Effects	Estimate	Std. Error	*t*-value	*p*MCMC
Intercept	3.018	.073	41.07	.0001
Coda /ŋ/	.011	.009	1.15	.237
Coda /s/	.029	.009	2.98	.003
Lexicality (real words)	-.014	.009	-1.46	.124
Cloze	.0009	.002	.57	.519
Coda /ŋ/ × Lexicality	-.018	.014	-1.35	.166
Coda /s/ × Lexicality	-.011	.014	-.83	.385

Korean

Response Times. In the model predicting response latency, Korean speakers only showed a significant effect of coda /s/ (Table 14). They were significantly faster to identify target words preceded by the coda /n/ compared to those preceded by the coda /s/ in nonword contexts. Pairwise comparison showed that this coda effect was not significant in the real word contexts ($z = 1.35$, $p = .756$)

Accuracy. No significant effect or interaction was found for the accuracy model (all *p*s > .1).

Table 14 Mixed-effects linear model estimating log RT in the Korean group

Fixed Effects	Estimate	Std. Error	*t*-value	*p*MCMC
Intercept	3.099	.040	76.65	.0001
Coda /ŋ/	.0003	.009	.03	.966
Coda /s/	.029	.009	2.95	.002
Lexicality (real words)	-.006	.009	-.58	.532
Cloze	.0002	.001	.2	.801
Coda /ŋ/ × Lexicality	-.009	.014	-.64	.492
Coda /s/ × Lexicality	-.016	.014	-1.17	.215

Mandarin

Table 15 Mixed-effects linear model estimating log RT in the Mandarin group

Fixed Effects	Estimate	Std. Error	*t*-value	*p*MCMC
Intercept	3.219	.052	62.36	.0001
Coda /ŋ/	.008	.009	.83	.368
Coda /s/	.029	.009	3.1	.002
Lexicality (real words)	-.004	.009	-.47	.6
Cloze	-.003	.001	-2.45	.011
Coda /ŋ/ × Lexicality	-.018	.013	-1.33	.159
Coda /s/ × Lexicality	-.014	.014	-1.01	.289

Response Times. Similar to the English and Korean groups, Mandarin speakers only showed a significant effect of coda /s/ in the model predicting response latency (Table 15). RTs were significantly longer when the target word was preceded by nonword contexts with coda /s/ compared to nonword contexts with coda /n/. This effect was not significant for real word context (*z* = 1.684, *p* = .542) based on pairwise comparison. There was a significant effect of cloze test scores, suggesting that higher proficiency is associated with faster segmentation of the target words. Pairwise comparisons did not show any additional significant findings (all *p*s > .1).

Table 16 Mixed-effects linear model estimating accuracy in the Mandarin group

Fixed Effects	Estimate	Std. Error	*t*-value	*p*MCMC
Intercept	4.582	1.317	3.479	.0005
Coda /ŋ/	-.285	.532	-.535	.592
Coda /s/	-1.015	.495	-2.052	.04
Lexicality (real words)	1.301	.771	1.687	.091
Cloze	-.029	.033	-.871	.384
Coda /ŋ/ × Lexicality	-.860	.947	-.908	.364
Coda /s/ × Lexicality	-.785	.894	-.878	.380

Accuracy. In the model predicting accuracy (Table 16), there was a significant effect of coda /s/ in which accuracy was lower when the target word had nonword contexts with coda /s/ compared to nonword contexts with coda /n/. There was a marginally significant effect of context lexicality in the /n/ condition in which Mandarin speakers' accuracy was higher when the target word was preceded by a real word context than by a nonword context. Pairwise comparisons did not show any additional significant findings (all *p*s > .1).

Spanish

Response Time. In the model predicting response latency (Table 17), Spanish speakers only showed a marginally significant effect of cloze test scores, suggesting higher proficiency is associated with faster segmentation time of the target words. Pairwise comparisons did not show any additional significant findings.

Accuracy. No significant effect or interaction was found in the accuracy model.

Table 17 Mixed-effects linear model estimating log RT in the Spanish group

Fixed Effects	Estimate	Std. Error	*t*-value	*p*MCMC
Intercept	3.218	.066	48.93	.0001
Coda /ŋ/	.0003	.011	.03	.959
Coda /s/	.016	.011	1.4	.14
Lexicality (real words)	-.014	.011	-1.25	.192
Cloze	-.003	.002	-1.68	.067
Coda /ŋ/ × Lexicality	-.007	.016	-.42	.65
Coda /s/ × Lexicality	.001	.016	.06	.952

Discussion

In summary, English, Korean, and Mandarin speakers all showed a significant effect of coda cue in which segmentation was slower or less accurate following coda /s/ compared to coda /n/. There was a significant lexicality effect for English speakers in the /ŋ/ condition. Spanish speakers did not show any significant effect for coda or lexicality.

Coda /s/

English, Korean, and Mandarin speakers experienced more difficulty segmenting target words preceded by /s/ coda than those preceded by /n/ coda. Given that both the native and nonnative groups show similar patterns of result, it is unlikely that the coda /s/ effect results from the influence of L1 phonotactic constraints. It would be incorrect to consider the coda /s/ effect as evidence for the use of phonotactic legality in segmentation since both /n/ and /s/ are legal in both syllable-initial and syllable-final positions in English. However, this could be considered the use of phonotactic probability in native English segmentation. Since /s/ has a similar likelihood of occurrence in word-initial and word-final positions (Dewey, 1950), particularly for multimorephemic words like the stimuli used in the current experiment, English listeners might take longer to decide whether to place a word boundary before or after /s/. In contrast, it was easier to identify a word boundary after /n/ since /n/ has a higher likelihood of occurrence in word-final than in word-initial position (Dewey, 1950; Vivevitch, & Luce, 2004).

There is another possible explanation for the coda /s/ effect which is the less robust allophonic cues of fricatives (Christie, 1974; Lehiste, 1960; Nakatani & Dukes, 1977). Allophonic variations are defined as the difference in acoustic details of how phonetic segments are pronounced in various syllable positions in fluent speech (Newman et al., 2011). Lehiste (1960) identified fricatives as the consonant class with weak allophonic cues. Only a few potential acoustic cues such as the duration of the fricative and the duration of the preceding vowel can differentiate between syllable-initial and syllable-final fricatives. In addition, these acoustic qualities are smaller and less consistent compared to those in voiceless stops (Christie, 1974; Nakatani & Dukes, 1977). Thus, it may be harder for native English speakers in the current study to identify word boundaries following coda /s/ as compared to coda /n/. Newman et al. (2011) found that consonants with stronger allophonic cues have a greater impact on segmentation than consonants with weaker allophonic variations. These researchers found that English speakers were faster to identify the embedded word if word boundary coincides with syllable boundary (i.e. *vuff-apple* easier to segment than *vuh-fapple*) and this effect was stronger for consonants with stronger allophonic cues. Newman et al. (2011) suggested that the strength of phonotactic probability as a segmentation cue is weaker than that of allophonic cues. Thus, even though /s/ is a legal coda in English, its weaker allophonic variations may make it more difficulty to use as a cue to word boundaries compared to /n/. This suggests that acoustic-phonetics influences native English segmentation. The design of the current experiment did not allow us to tease apart whether acoustic-phonetics or probability phonotactics influenced English speakers' segmentation of coda /s/. Since both of these cues are segmental cues, the only safe conclusion is that English speakers did use segmental cues in native segmentation.

The acoustic qualities, phonotactic probability, or a combination of both of coda /s/ may have also influenced Korean and Mandarin speakers' segmentation of L2 words assuming they have acquired sensitivity to the phonotactics and acoustic-phonetics in English. Alternatively, Korean and Mandarin speakers' difficulty with /s/ coda may be partially influenced by the phonotactic constraints in their L1. The latter hypothesis would be supported if there is a significant language × coda interaction when comparing the native and L2 groups and comparing /n/ and /s/. In a generalized mixed-effects linear model estimating reactions times with the English group as the baseline, language group, coda, and cloze test scores were entered as fixed effects while subject and items were entered as random effects. The interaction between language group and coda was not significant (Korean vs. English: $t = -.51$, $p = .606$; Mandarin vs. English: $t = -.10$, $p = .907$). These results suggest that the fact that /s/ violates L1 phonotactic constraints did not increase Korean and Mandarin speakers' difficulty with /s/ coda in L2 segmentation. It appears that L2 learners were influenced by acoustic-phonetics or phonotactic probabilities in English to the same degree as the native speakers.

Coda /ŋ/

None of the language groups showed a significant effect of coda /ŋ/. The absence of this coda effect in English, Korean, and Mandarin speakers is not consistent with the hypotheses. Since /ŋ/ is only allowed syllable-finally in all three languages, it should be easier for listeners to identify a word boundary following /ŋ/ than following /n/. Given that both native and nonnative groups showed similar patterns of result with regards to /ŋ/, it is unlikely that L1 phonotactic constraints play a role in the use of this coda cue in L2 segmentation.

It is surprising that Spanish speakers did not show more difficulty segmenting words with coda /ŋ/ even though this phoneme does not exist in their L1. It is possible that the absence of this phoneme in L1 may allow Spanish speakers to establish a new phonological category in the process of L2 acquisition and potentially use this phoneme to segment L2 speech. The Speech Learning Model (SLM) (Flege, 1988, 1995, 1999, 2002) proposes that the more dissimilar L2 sounds are from the closest native phonemes, the better likelihood of establishing new L2 phonological categories. Results from Experiment 1 showed that the absence of a phonological feature in L1 does not necessarily prevent nonnative speakers from using this phonological feature (in a way that is consistent with L1 typology) to segment L2 speech (i.e. stress cues for Korean speakers). Thus, it is possible that Spanish speakers have established a phonological category for /ŋ/, which allows them to learn the phonotactic probability of this coda cue through repeated L2 exposure or classroom instruction.

Lexical Cues

None of the L2 groups showed a significant lexicality effect in any of the three coda conditions. However, the significant coda /s/ effect observed in Korean and Mandarin speakers are only present in nonword context. The absence of the coda /s/ effect in real word context suggests that lexical knowledge may help compensate for some of the difficulties segmenting words with coda /s/. In other words, Korean and

Mandarin speakers may be less likely to consider /s/ as the onset of the following word or less influenced by the weaker acoustic qualities of fricatives if they recognize the preceding word. Given the interaction between coda /s/ and lexicality was not significant in both language groups, it appeared that acoustic-phonetics or probabilistic phonotactics are relatively more important than lexical knowledge.

English speakers only showed a significant lexicality effect in the /ŋ/ condition but not in the /n/ or /s/ condition. This result may be explained by conceptualizing the use of lexical cues as a function of phonotactic probability (rather than absolute legality). All three phonemes are perfectly legal in word-final positions, yet, they differ in phonotactic probabilities. The probability of /ŋ/ in word-initial position is 0%. In comparison, the probability of /s/ or /n/ in word-initial positions is more than 0% as they are allowed both word-initially and word-finally. Thus, it is possible that a single sound with 0% phonotactic probability at word edges facilitates the use of lexical knowledge. This reasoning would suggest that native speakers, like the nonnative speakers, rely more on probabilistic phonotactics than lexical cues. Since the current experiment did not examine the effect of phonotactic probability on the use of phonotactic cues in segmentation, this explanation needs to be tested empirically in future research. If this explanation holds true, then segments with 0% phonotactic probability at the word-final position such as the consonant cluster /sw/ would also facilitate the use of lexical knowledge in native English segmentation.

Chapter 6 - Experiment 3 Phonotactic Cues versus Semantic Cues

This experiment compared the use of phonotactic cues and semantic cues in L2 segmentation. Experiment 3 was largely similar to Experiment 2 in terms of using the same codas (e.g., /n, ŋ, s/) to examine the use of phonotactic cues. Experiment 3 also employed a similar word spotting task. There were two main differences. First, only real word stimuli were used in this experiment since lexical knowledge is not a variable of interest here. Second, the use of semantic cues was examined by manipulating the semantic relatedness of *auditory context* and *auditory target*. Previous research has shown that participants react faster to the target word *nurse* if they have heard *doctor* before (Perea & Rosa, 2002).

Materials and Design

The design was 2 × 3 factorial with the two factors being the semantic relatedness of the *context* and *target* (related or unrelated) and syllable coda in the *context* (/n, ŋ, s/). In the critical trials, all *auditory contexts* were trisyllabic words whereas all *auditory targets* were disyllabic. Twenty disyllabic high-frequency nouns with concrete meanings were generated from the English Lexicon Project (Balota et al., 2007). For each target word, three semantically related words and three semantically unrelated words were created for each of the three coda conditions. Thus, the original set of stimuli consists of a total of 360 context-target phrases. This list was sent to the L2 raters who completed the familiarity ratings for the previous experiments and an additional 10 native English speakers. They were asked to judge how much is the target word related to the context on a 7-point Likert scale with 1 being “very unrelated” and 7 being “very related”. The list was pseudo-randomized so that the raters did not see the same target words consecutively.

For each coda condition, the context word with the highest relatedness rating was chosen for the related condition and the context word with the lowest relatedness rating was selected for the unrelated condition. It was also ensured in the /s/ coda condition that the word-final /s/ is not realized as /z/ in all of the words. This resulted in 20 related and 20 unrelated context words for each coda condition. Context words in the related condition were rated significantly more related to the target words than those in the unrelated condition ($F(118) = 21.12$, $p < .001$). Words in the related and unrelated conditions were matched on all of the relevant properties as in previous experiments (Table 18). Words in the three coda conditions were also well matched.

Table 18 Mean Statistics of the stimuli from Experiment 3

	Coda					
	n		ŋ		s	
	Relate	Unrelated	Relate	Unrelated	Relate	Unrelated
Semantic relatedness (out of 7)	5.86	1.96	5.88	1.99	5.14	1.49
Written frequency (log)	8.44	8.31	7.26	7.64	8.36	7.89
Spoken frequency (log)	2.35	2.31	2.12	2.01	2.27	2.17
Number of letters	8.5	8.25	9.3	8.9	9.35	9.2
Number of phonemes	7.15	6.6	7.4	7.15	8.0	7.65
Biphone frequency (token)	220.83	175.13	651.29	798.56	241.86	214.82
Biphone frequency (type)	19.73	17.41	117.64	124.31	25.24	26.6
Phonological neighborhood size	.5	.7	1.1	1.4	.2	.5
Uniqueness point	3.1	3.25	3.05	2.95	3.3	3.45

As a result of the 2 × 3 factorial design, each target was paired with one semantically related and one semantically unrelated context in each of the three coda conditions, yielding six possible combinations. To prevent repeated exposure to the same target, three phrases were selected from each set of six and the selection was counterbalanced based on a Latin-square design. Two lists were created so that List 1 contained three phrases from one set and List 2 contained the other three phrases from the same set. There were a total of 60 critical trials in each list. Appendix C shows a complete list of the stimuli from Experiment 3.

To prevent participants from developing processing strategies for words ending with /n ŋ s/ or only focusing on the last three syllables of the auditory phrase, three types of fillers were created, with 60 trials in each type. For the first type of filler trials, the visual targets were disyllabic words that matched neither the auditory contexts nor the auditory targets. For the second type of filler trials, the visual targets were trisyllabic words that matched the auditory contexts. Finally, for the third type of filler trials, the visual targets were also trisylalbic words but they matched neither the auditory contexts nor the auditory targets. There were a total of 240 trials, 60 critical trials and 180 filler trials, with an equal number of positive and negative responses. The trials were pseudo-randomized so that there were at least 70 trials separating the same context to prevent any repetition priming effect. There were no more than three "Yes" or "No" responses consecutively. The same female native English speaker recorded all stimuli using the same methods as in previous experiments.

Procedure

Procedure was the same as that in Experiment 2.

Hypotheses

When syllable codas are consistent with L1 phonotactic constraints (e.g., /s/ for Spanish speakers and /ŋ/ for Mandarin and Korean speakers), participants were expected to respond faster to words preceded by semantically related contexts than those preceded by unrelated contexts (Table 19); when syllable codas do not conform to L1 phonotactic constraints (e.g., /s/ for Mandarin and Korean speakers and /ŋ/ for Spanish speakers), the semantic relatedness effect were hypothesized to be smaller. It is probable that more advanced L2 learners would utilize semantic information more than phonotactic cues. However, it remains unclear what is the threshold of proficiency necessary for L2 learners to utilize semantic cues in segmentation. During stimuli selection, it was ensured that only familiar words with clear semantic relations were chosen for this experiment. It would be reasonable to expect that L2 learners could use semantic cues, but this usage may not be as automatic and efficient as that in native speakers. Consequently, L2 learners might still rely more on sublexical cues.

Table 19 Predictions (larger numeric magnitude represents greater weight given to that cue)

	English	Mandarin	Korean	Spanish
Phonotactic cues	1	2	2	2
Semantic cues	2	1	1	1

Data Analyses

Data from 31 Korean, 22 Spanish, 30 Mandarin, and 35 English speakers were analyzed. The dependent variables were accuracy and (log-transformed) RT in the word spotting task. RT data for incorrect responses were excluded (5.3% of total data). The independent variables were the coda of the auditory context and semantic relatedness between the context and target. The analytical method was similar to that in Experiment 2. The coda /n/ condition was treated as the intercept in the linear mixed-effects model for each language group. Coda and semantic relatedness, and cloze test scores (as a covariate for proficiency) were the fixed effects while participant and item were the random effects (results from the model with language group as an additional fixed effect is shown on

Appendix G). Accuracy data was analyzed using the same mixed-effects model but with the binomial function. For brevity, only statistically significant results were reported.

Results

Table 20 shows the mean latencies and accuracy rates in the word spotting task. The main findings of this experiment (Figures 8 and 9) were that all four language groups showed an effect of coda cue in which they responded slower or less accurate to target words preceded by /s/ than those preceded by /n/. Korean speakers showed a significant two-way interaction between semantic relatedness and coda cues in which the semantic effect was present in the /n/ condition but absent in the other two coda cues. English speakers also showed a significant semantic effect in the /n/ condition but semantic relatedness did not interact with coda cues.

Table 20 Mean word spotting RTs and % accuracy in parentheses

	English		Korean		Mandarin		Spanish	
Semantic	Relate	Unrelated	Related	Unrelated	Related	Unrelated	Related	Unrelated
Coda /n/	1109 (97)	1138 (94)	1233 (95)	1312 (93)	1216 (97)	1229 (98)	1207 (98)	1221 (97)
Coda /ŋ/	1089 (97)	1096 (95)	1222 (94)	1224 (96)	1194 (97)	1202 (97)	1187 (98)	1201 (97)
Coda /s/	1176 (95)	1189 (92)	1361 (89)	1351 (89)	1319 (93)	1308 (92)	1332 (92)	1314 (89)

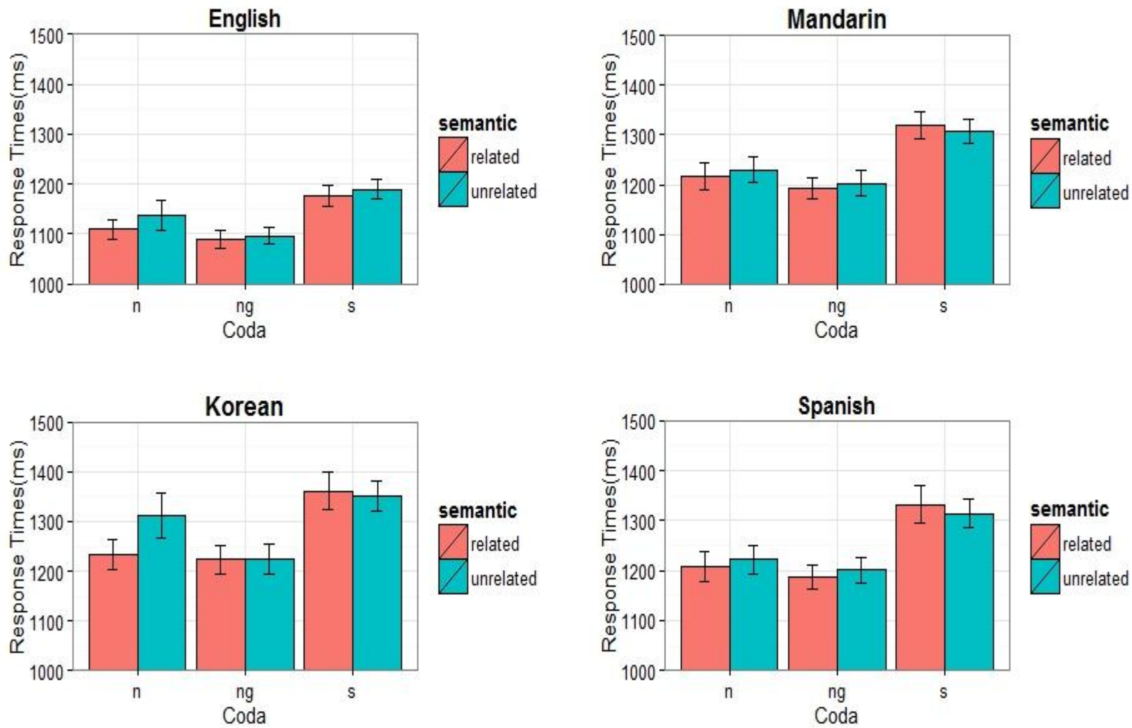

Figure 8 Phonotactic Cues vs. Semantic Relatedness (RT)

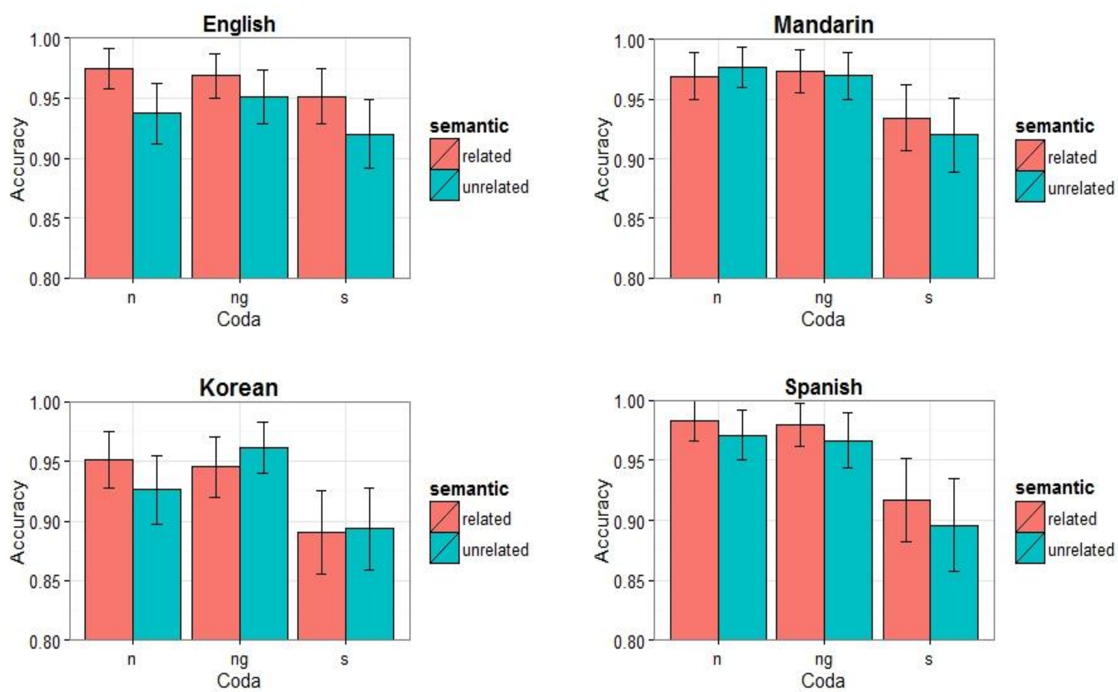

Figure 9 Phonotactic Cues vs. Semantic Relatedness (Accuracy)

English

Response Time. In the model predicting response latency (Table 21), the intercept estimated English speakers' log RT when the target words were preceded by a semantically related context with coda /n/. English speakers showed a significant effect of coda /s/ for semantically related contexts. They speakers were significantly faster to identify target words preceded by coda /n/ compared to those preceded by /s/. This was also true for semantically unrelated contexts ($z = 4.799$, $p < .001$) based on the pairwise comparison. There was a marginally significant effect of semantic relatedness for coda /n/. English speakers were faster to spot target words preceded by semantically relatedness contexts than those preceded by semantically unrelated contexts. Pairwise comparisons showed no significant difference between related and unrelated conditions for both /ŋ/ and /s/ (all *p*s > .l).

Table 21 Mixed-effects linear model estimating log RT in the English group

Fixed Effects	Estimate	Std. Error	*t*-value	*p*MCMC
Intercept	3.082	.063	48.73	.0001
Coda /ŋ/	-.007	.004	-1.46	.145
Coda /s/	.026	.004	5.71	.0001
Semantic relatedness (unrelated)	.009	.004	1.93	.054
Cloze	-.001	.001	-.55	.544
Coda /ŋ/ × Semantic relatedness	-.005	.006	-.83	.406
Coda /s/ × Semantic relatedness	-.004	.006	-.59	.554

Accuracy. In the model predicting response accuracy, there was only a significant effect of semantic relatedness in the coda /n/ condition ($z = -2.285$, $p = .022$). Word identification was significantly more accurate when the target word was preceded by the semantically related context than those preceded by unrelated context. All other effects and interactions were not significant (all *p*s > .1).

Korean

Response Time. In the model predicting response latency (Table 22), the intercept estimated Korean speakers' log RT when the target words were preceded by a semantically related context with coda /n/. There was a significant interaction between semantic relatedness and codas /n- ŋ/. Korean speakers showed a significant semantic effect in the /n/ condition in which they were faster to identify target words preceded by semantic related contexts than those preceded by semantic unrelated contexts (z = 4.05, p < .001). However, for both /ŋ/ and /s/, RT did not significantly differ between target words preceded by semantically related context and those preceded by unrelated context (both ps = .999). For semantically related contexts, Korean speakers showed significantly faster RT for target words in the /n/ condition than those in the /s/ condition. Pairwise comparisons showed that, for semantically unrelated contexts, Korean speakers also showed significantly faster RT for target words preceded by /ŋ/ (z = -4.444, p < .001) and significantly slower RT for target words preceded by /s/ (z = 3.188, p = .017) than those preceded by /n/.

Table 22 Mixed-effects linear model estimating log RT in the Korean group

Fixed Effects	Estimate	Std. Error	*t*-value	*p*MCMC
Intercept	3.062	.038	81.06	.0001
Coda /ŋ/	-.0002	.006	-.03	.956
Coda /s/	.041	.006	6.97	.0001
Semantic relatedness (unrelated)	.024	.006	4.05	.0001
Cloze	.001	.001	.55	.523
Coda /ŋ/ × Semantic relatedness	-.026	.008	-3.12	.003
Coda /s/ × Semantic relatedness	-.022	.008	-2.67	.008

Accuracy. In the model predicting accuracy, there was only one significant effect of coda /s/ in semantically related contexts (z = -2.722, p = .006). Korean speakers were significantly less accurate identifying target words preceded by coda /s/ than those preceded by /n/. This semantic effect was also present for semantic unrelated contexts since the two-way interaction between coda and semantic relatedness was not significant (z = 1.057, p = .291.)

Mandarin

***Response Times*.** In the model predicting response latency (Table 23), the intercept estimated Mandarin speakers' log RT in the coda /n/ condition with semantically related contexts. For semantically related contexts, there was a significant effect of /s/. Compared to target words preceded by coda /n/, Mandarin speakers were significantly slower to identify those preceded by /s/. Pairwise comparison showed that this effect was also present for semantically unrelated context (z = 5.785, p < .001). There was a marginally significant effect of /ŋ/ in which Mandarin speakers were faster to segment target words preceded by /ŋ/ than those preceded by /n/ in semantically related contexts. Pairwise comparisons did not show any other significant effect (all *p*s > .1). There was a significant effect of cloze test in which higher cloze test scores were associated with faster RT in the word-spotting task.

Table 23 Mixed-effects linear model estimating log RT in the Mandarin group

Fixed Effects	Estimate	Std. Error	*t*-value	*p*MCMC
Intercept	3.238	.055	59.01	.0001
Coda /ŋ/	-.009	.005	-1.79	.072
Coda /s/	.035	.005	7.29	.0001
Semantic relatedness (unrelated)	.004	.005	.92	.351
Cloze	-.004	.001	-2.92	.004
Coda /ŋ/ × Semantic relatedness	-.0001	.007	-.02	.983
Coda /s/ × Semantic relatedness	-.007	.007	-1.08	.262

Accuracy. In the model predicting accuracy, there was no significant effect or interaction (all *p*s > .05). Pairwise comparisons showed that Mandarin speakers were significantly less accurate to identify target words with coda /s/ than those with coda /n/ for semantically unrelated context ($z = -3.073$, $p = .025$). However, this effect was not significant for semantically related contexts ($z = -1.852$, $p = .424$).

Spanish

Response Times. In the model predicting response latency (Table 24), the intercept estimated Spanish speakers' log RT for coda /n/ and semantically related contexts. There was a significant effect for coda /s/. For semantically related contexts, Spanish speakers were significantly slower to identify target words preceded by /s/ than those preceded by /n/. Pairwise comparisons showed the same pattern of results for semantically unrelated contexts ($z = 5.339$, $p < .001$).

Table 24 Mixed-effects linear model estimating log RT in the Spanish group

Fixed Effects	Estimate	Std. Error	*t*-value	*p*MCMC
Intercept	3.14	.056	56.51	.0001
Coda /ŋ/	-.004	.006	-.77	.444
Coda /s/	.041	.006	7.33	.0001
Semantic relatedness (unrelated)	.006	.006	1.06	.289
Cloze	-.002	.001	-1.19	.179
Coda /ŋ/ × Semantic relatedness	-.002	.008	-.26	.788
Coda /s/ × Semantic relatedness	-.011	.008	-1.39	.167

Accuracy. In the model predicting accuracy, there was a significant effect of coda /s/ for semantically related contexts ($z = -2.826$, $p = .005$). Identification accuracy was significantly lower for target words preceded by /s/ than those preceded by /n/. This coda effect was also significant for semantically unrelated contexts ($z = -3.253$, $p = .013$) based on pairwise comparison.

Comparing Semantic Relatedness and Lexical Knowledge

Post-hoc analyses were conducted to compare the relative importance of semantic cues and lexical knowledge in L2 segmentation by combining the datasets from Experiments 2 and 3. In generalized linear mixed-effects models, cue type (two levels: semantic or lexical), coda (three levels: n, ŋ, s), and cloze test scores were entered as fixed effects and participants and items were entered as random effects. For brevity, only significant results from RT analyses were reported (Figure 10).

English. When cue type was lexical knowledge, English speakers were significantly faster to segment context words with coda /ŋ/ (t = -2.02, p =.045) but slower to segment context words with coda /s/ (t = 2.09, p = .039) compared to context words with coda /n/. The two-way interaction between cue type and coda (/n- ŋ/ and /n-s/) was not significant (both ps > .1). Thus it can be inferred that the coda effects were also present when cue type was semantic relatedness.

Korean. There was a significant effect of cue type for coda /n/. Korean speakers were significantly faster to segment context words that are semantically related to the target words than those that are real words but semantically unrelated (t = -3.77, p =.0001). There was a two-way interaction between cue type and coda /n-s/. When the cue type was lexical knowledge, RT did not differ significantly between coda /n-s/ (t = 1.26, p =.211); but when the cue type was semantic relatedness, Korean speakers were significantly faster to segment context words with coda /n/ than those with coda /s/ (t = 3.05, p =.002).

Mandarin. Similar to Korean speakers, Mandarin speakers showed a significant two-way interaction between cue type and coda /n-s/. When cue type was lexical knowledge, RT was marginally significantly faster for context words with coda /n/ than those with coda /s/ (t = 1.92, p =.056); when cue type was semantic relatedness, the RT difference between /n-s/ was significantly larger (t = 2.00, p =.046). There was also a significant cue type effect for coda /n/ (t = -2.22, p =.027). Mandarin speakers were significantly faster to segment context words that were semantically related to the following target words than those that are real words but semantically unrelated.

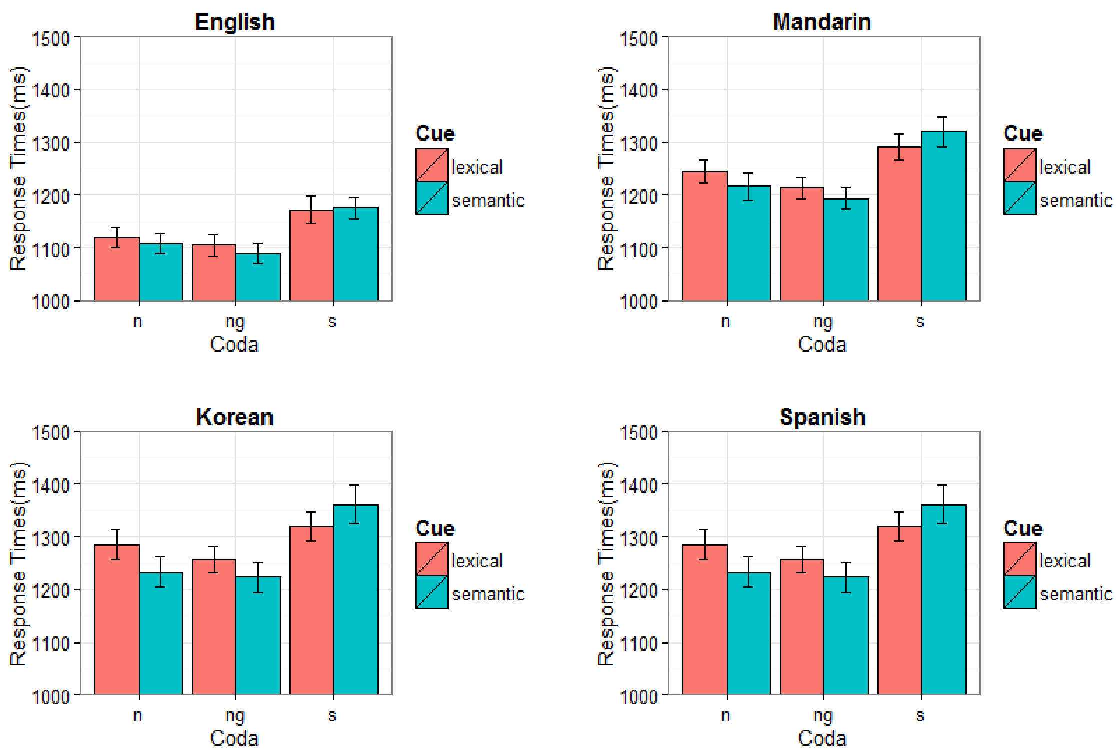

Figure 10 Comparing lexical knowledge and semantic relatedness (RT)

Spanish. Similar to both Mandarin and Korean speakers, there was a significant two-way interaction between cue type and coda /n-s/. When the cue type was lexical knowledge, RT was marginally significantly faster in the /n/ condition compared to the /s/ condition (t = 1.95, p =.055); but when the cue type was semantic relatedness, the RT difference was significantly larger (t = 2.13, p =.032). Spanish speakers were also significantly faster to segment context words with coda /n/ that are semantically related to the target words than those that are real words (but semantically unrelated to the target words) (t = -2.53, p =.009).

Discussion

Overall, the results with the coda cues are similar to those in Experiment 2. English, Mandarin, and Spanish speakers all showed a significant coda effect in which they were significantly slower or less accurate to identify target words preceded by /s/ than those preceded by /n/ for both related and unrelated contexts. English speakers also showed a marginally significant semantic effect in the /n/ condition. Korean speakers were the only language group to show evidence of cue interaction. Korean speakers were slower to identify target words preceded by /s/ compared to those preceded by /n/ for related contexts whereas their response times did not differ significantly between /n-s/ for unrelated contexts.

Coda /s/

English, Mandarin, and Spanish speakers experienced more difficulty segmenting words with coda /s/ than those with coda /n/ and these findings are similar to those in Experiment 2. Since both native and nonnative speakers showed similar response patterns, the coda /s/ effect should not be considered as an effect of L1 transfer. As previously mentioned, /s/ has weaker allophonic cues (Lehiste, 1960) so that the acoustic qualities used to distinguish between word-initial and word-final fricatives are smaller and less consistent. Although word-final /s/ or /n/ conforms to the phonotactic constrains in English, it is possible that the weaker allophonic variation makes it more difficult to segment words with coda /s/ than those with coda /n/ due to allophonic cues having greater strength than phonotactic cues in English segmentation (Newman et al., 2011). Also, the relative frequency of /s/ in word-initial and word-final positions is similar (Dewey, 1950). Thus, English listeners might take longer to decide whether to group /s/ with the following or preceding strings. The current experiment could not tease apart the two possibilities. Since both acoustic-phonetics and probabilistic phonotactics are segmental cues, the only safe conclusion is that English speakers utilized segmental cues in native segmentation. In addition, the fact that English speakers showed a significant effect of coda cue but only a marginally significant effect of semantic relatedness suggests that, in the context of the current study, English speakers rely more segmental cues than word-level semantics.

Like native English speakers, Mandarin and Spanish speakers may have been influenced by the less salient acoustic properties of fricatives compared to other consonant classes such as nasals and stops (Christie, 1970; Nakatani & Dukes, 1977) or the similar relative frequency of /s/ as an onset or a coda in English (Dewey, 1950). As a result, the nonnative listeners took longer to identify word boundaries following /s/ than following /n/. However, /s/ in the word-final position violates the phonotactic constraints in Mandarin and word-final /s/ has numerous phonological variations in different Spanish dialects. These characteristics of /s/ in L1 may have additive effect on Mandarin and Spanish speakers' difficulty in L2 segmentation. This possibility was explored in a generalized linear mixed-effects model estimating response times with the English group as the baseline. Language group, coda cues (/n / vs. /s/), and cloze test scores were entered as fixed effects while subject and items were entered as random effects. The two-way interaction between language group and coda cue was significant in the Spanish and

English comparison ($t = 2.43$, $p = .013$) but not in the Mandarin and English comparison ($t = 1.57$, $p = .119$). In other words, native English speakers took significantly longer to segment words with coda /s/ than those with coda /n/ and this difference in response times is significantly larger in Spanish speakers. These results suggest that, in addition to the influence of weaker allophonic variation or phonotactic probabilities, the violation of L1 phonotactic constraints did not add to Mandarin speakers' difficulty with segmenting words with coda /s/ whereas the phonological variation of word-final /s/ in L1 adds to Spanish speakers' difficulty. This result suggests that there is a possibility that L1 acoustic-phonetics influences the use of cues in L2 segmentation.

In both a lexical decision task and a naming task of Spanish words, Bommershine (2006) found that both native speakers of Puerto Rican Spanish and native speakers of Mexican Spanish were significantly slower to respond to words with coda /s/ than those with coda /n/. It appears that the phonological variation of word-final /s/ has a greater impact on the processing of L1 words compared to word-final /n/. Participants in the Spanish group of the current dissertation project came from a variety of Spanish-speaking countries including Mexico, Spain, Chile, Colombia and Puerto Rico. Due to the heterogeneity of the language background of the Spanish speakers, it would not be useful to analyze Spanish speakers' performance base on L1 dialects. However, findings of the current experiment appear to suggest that, Spanish speakers might have more difficulty segmenting L2 words with /s/ coda compared to those with /n/ coda due to the influence of phonological variation of /s/ in L1.

Coda /ŋ/

Korean speakers showed a significant effect of coda /ŋ/ for semantically unrelated context. Since /ŋ/ can only occur syllable-finally in Korean, it may be easier for Korean speakers to predict a word boundary and recognize the target words when they hear /ŋ/ in the context words. The coda /ŋ/ effect was not significant for semantically related contexts. The significant interaction between coda /ŋ/ and semantic relatedness suggests that the segmentation advantage of coda /ŋ/ may be minimized when Korean speakers could utilize semantic cues to identify word boundaries (see section "Cue Interaction" below). English, Mandarin, or Spanish speakers did not show any significant effect of coda /ŋ/ and these findings are not consistent with the hypotheses. English and Mandarin speakers were expected to be faster to segment words with coda /ŋ/ is a legal coda in English and Mandarin. Spanish speakers were expected to have difficulty segmenting words with coda /ŋ/ since this phoneme does not exist in their L1. These findings replicated those in Experiment 2, suggesting that there was no L1 influence on the use of this coda cue in the Mandarin and Spanish groups. Particularly for Spanish speakers, their native-like behavioral responses of /ŋ/ suggest that they might have developed a phonological category for this phoneme and were able to segment words containing /ŋ/ efficiently

Semantic Cues

English speakers showed evidence of using semantic relatedness as a segmentation cue in the /n/ condition only. They were faster and more accurate to identify the target words preceded by semantically related contexts than those preceded by unrelated contexts. Semantic relatedness did not interact with phonotactic cues and this seems to suggest that English speakers use the cues independently. Mandarin and Spanish speakers did not show a significant effect of semantic relatedness in any of the three coda conditions. Since both groups showed a significant effect of coda /s/, it appears that both L2 groups rely more on acoustic-phonetic cues than semantic relatedness.

Cue Interaction

Korean speakers are the only language group to show a significant interaction between phonotactic cues and semantic cues. For semantic related context, Korean speakers were significantly faster to recognize target words with coda /n/ than those with coda /s/; for semantically unrelated context, the difference in response times between the two coda conditions were significantly smaller. These results suggest that identifying a word boundary following coda /s/ is more difficult when semantic cues are available. This finding is a direct opposite of that for coda /ŋ/ in which segmentation following coda /ŋ/ is easier when semantic cues are not available. It appears that when coda cue is consistent with L1 phonotactic constraint (i.e. word-final /ŋ/), semantic relatedness did not facilitate segmentation; when coda cue violates L1 phonotactic constraint (i.e. word-final /s/), semantic relatedness makes the identification of word boundaries faster. Although Mandarin and Korean speakers have similar phonotactic constraints regarding /s/ and /ŋ/ in their respective L1s, the two language groups showed vastly different use of phonotactic and semantic cues. Mandarin speakers appeared to rely more on acoustic-phonetic cues whereas Korean speakers used both cues interactively. One striking similarity of the L2 groups is that their use of cues seemed to be influenced by L1 typology (see General Discussion).

Cue Type: Lexical Knowledge vs. Semantic Relatedness (Experiments 2 and 3)

Comparing the data from Experiments 2 and 3 showed that all three L2 groups were faster to segment context words with coda /n/ that were semantically related to the following target words (e.g., *syllabus mentor*) than context words that were real words but semantically unrelated to the target words (e.g., *evidence already*). In addition, all three L2 groups were slower to segment context words with coda /s/ compared to those with coda /n/ and this difference in RT was significantly larger for semantic cues than for lexical cues. Since the stimuli used in Experiments 2 and 3 were different and the two types of cues were not directly tested in the same experiment, one must be cautious when drawing conclusion from these results. It appears that semantically related context facilitated segmentation to a greater degree than real word contexts that are not semantically related. In other words, knowing what constitutes a real word may not be as helpful as understanding its meaning in L2 segmentation. This finding is not consistent with what was hypothesized by RFL2. On the other hand, English speakers did not show any significant RT difference in cue type, suggesting that they did not necessarily rely more on word-level semantics than lexical knowledge. This finding is not consistent with the Hierarchical Framework (Mattys et al., 2005), which proposed that semantics are weighed over lexical knowledge in native English speakers' L1 segmentation. See General Discussion for more detailed discussion of this issue.

Chapter 7 - Experiment 4 Stress Cues versus Sentence Context

This experiment compared the use of stress cues and sentence context in L2 segmentation. The cross-modal priming task was adapted from Mattys et al. (2005, Exp.6). One novel difference of the current experiment is that SW words were added to the stimuli. In this experiment, participants heard sentences and saw a sequence of letter strings 100ms after the offset of the auditory presentation. They determined whether the letter strings represent a real English word. The first syllable in the SW targets (e.g., *CANcel*) is itself a real English word (e.g., *can*) while the second syllable in the WS targets (e.g., *enSURE*) is itself a real English word (e.g., *sure*). In the related sentence context condition, the semantic context of the spoken sentence encourages the segmentation of the disyllabic word but not the monosyllabic word within the target. In the unrelated condition, the semantic context of the sentence favors neither the disyllabic nor the monosyllabic parsing of the target word. Since the sentences in the related condition are predictable of the meaning of the disyllabic words, if speakers weigh sentence context over stress cues, word recognition time would be similar regardless of the stress patterns of the target words. Yet, if speakers assign more weight to stress cues over sentence context, the SW words would be recognized faster than the WS words since participants would erroneously segmented the S monosyllabic words in the WS targets at first parsing.

Materials and Design

The design is a 3 × 2 factorial design with the factors being semantic relatedness of the sentence context to the target words (baseline, related, or unrelated) and stress patterns of the target (e.g., SW or WS). An initial 30 pairs of disyllabic target words with high frequency and concrete meanings were generated from the English Lexicon Project (Balota et al., 2007). There was one trochee and one iamb in each pair. The S syllable in each word also is a real word in English (e.g., *form* in *perform*). Predictable and unpredictable sentence context was made for each target word. Two rounds of sentence predictability ratings were collected. In the first round, four different distracter words were created for each sentence context. Along with the target word, a total of five choices were available for each sentence context. The same raters who completed the ratings for previous experiments also did the sentence predictability ratings. Each sentence context was presented up to the point of the target word and the raters did not see the end of the sentence. Five choices were presented following the sentence fragment and the raters were asked to judge the likelihood of occurrence for each word based on the meaning of the sentence. The ratings were based on a 7-point Likert scale with 1 being "very unlikely" and 7 being "very likely". For example, participants read the sentence fragment "*The captain finally surrenders, seeing that enemies*____," followed by the five choices *attack*, *strike*, *surround*, *assault*, and *bomb*. The raters were informed that the target words do not necessarily have to finish the sentences. The order of the target words were randomized within each sentence and the order of the sentences were randomized across raters.

Among the five choices, if the target word received the highest likelihood rating, then the sentence was retained for the predictable condition. If the target word received the lowest likelihood rating, then the corresponding sentence was retained for the unpredictable condition. If one target word (e.g., a trochee) was deleted, the other target word (e.g., an iamb) in the pair was also deleted. After this first round of selection, only

22 pairs of target words and sentence contexts remained. In the second round of selection, the same set of word choices were created for both the related and unrelated sentence contexts. The five word choices included the target word, three distracters that were pragmatically and grammatically possible in both sentence contexts, and the word that has received the highest rating for the unrelated sentence context from the first round of ratings. For example, for the target word *cancel*, the unrelated context is "*An important client has arrived, the lawyers* ____," the distracter word with the highest likelihood rating from the first round was *welcome*. For the second round of ratings, the choices for this sentence context were *cancel*, *welcome, reschedule*, *arrange,* and *postpone*. Ten native English speakers who did not participate in any previous rating tasks completed the second round of ratings. As in the first sound, if the target word received the highest rating, the corresponding sentence was kept for the predictable condition. If the target word did not get the highest rating, the sentence was retained for the unpredictable condition. An additional two pairs of sentence contexts were deleted as a result of this procedure. The final stimuli set consisted of 20 pairs of disyllabic words, half of them were trochees and half of them were iambs. Each disyllabic target word had a predictable sentence context and an unpredictable context. Appendix D shows a complete list of stimuli in Experiment 4. Forty monosyllabic S words from the disyllabic words also served as target words to encourage participants to consider the monosyllabic words for segmentation. The disyllabic and monosyllabic words were matched on written frequency, spoken frequency, and biphone frequency (Table 25). The SW and WS words were matched on all the relevant properties.

Table 25 Mean Statistics for Stimuli in Experiment 4

	Target Words		
	Monosyllabic	Disyllabic	
		SW	WS
Written frequency (log)	10.29	10.67	9.96
Spoken frequency (log)	3.44	3.15	3.22
Phonological neighborhood size	22.8	4.3	1.3
Biphone frequency	1071.99	1033.08	512.85
Number of letters	3.95	6.1	6.7
Number of phonemes	3.45	5.1	5.7
Uniqueness point	4.95	7.1	7.7

In both the related and unrelated sentence contexts, auditory target words were always immediately followed by their visual presentation. Baseline sentences matched in duration with the critical sentences were created to assess the phonological repetition priming effect. A total of 40 baseline sentences that did not contain any phonological overlap with the target words were created. Two test lists were created and half of the critical sentences were paired with monosyllabic target words and half of them were paired with disyllabic target words. For example, in List 1, the related sentence context was paired with the target word *cancel* and the baseline sentence matched in duration was paired with the target word *can*. Similarly in List 2, the unrelated sentence context was paired with *can* while the baseline sentence was paired with *cancel*. Each target word did not appear more than once in each list. The two factors (sentence context and stress patterns) were counterbalanced using a Latin-square design across the two lists. There were a total of 80 critical trials in each list.

Two types of filler trials were created. For the first type of filler trials, the target word were semantically related to the preceding spoken word in the sentence and half of the target words were monosyllabic and the other half were disyllabic. For example, the sentence context "*He worked hard for many companies to further his career*____" was paired with the target word *job*. For the second type of filler trials, the target words were all nonwords and they overlapped with the preceding spoken word in the sentence. The portion of overlap is full, partial, or none, with 40 trials of each type. For example, the sentence context "*To solve the problem, a reward was offered to inventors who could create the device*____" was paired with the visual target (i.e. two syllable letter strings) *thevice* which fully overlapped with the preceding spoken word. There were 40 trials in the first type of fillers and 80 trials in the second type. Each test list had a total of 160 trials, with an equal number of positive and negative responses. The trials were pseudo-randomized so that there were at least 70 trials separating the disyllabic target word (e.g., *perform*) and its corresponding monosyllabic word (e.g., *form*). It was also ensured that there was no more than three "Yes" or "No" responses consecutively.

The same female native English speakers recorded all the stimuli using the same method as in previous experiments. She recorded all the sentences continuously. Each sentence was manually cut into two parts at the boundary between the prime word and the following word (In Appendix D). To ensure that participants not only focus on the visual targets but also pay attention to the sentence context, 80 true or false comprehension questions were created for half of the 160 sentences in each list. Participants had to answer the question based on the context of the sentence. Half of the 80 questions had a "Yes" response and the other half had a "No" response.

Procedure

Participants were tested individually using a desktop or laptop PC in a quiet room. They were randomly assigned to List 1 or List 2. The experiment was implemented via the E-prime software (Psychology Software Inc., Pittsburgh, PA) which automatically collected response latency and accuracy. In each trial, participants first saw a fixation sign "+" in the center of the screen for 500ms. Then they heard the sentence fragment, which is followed by the visual presentation of letter strings, written in 22 pt bold Courier font, 100ms after the offset of the auditory sentence. Participants were instructed to decide whether the letter strings represent a real English word by pressing the keys labeled "Yes" or "No" on the computer keyboard. After the lexical decision is made, there was an inter-stimulus interval of 500ms. Then participants heard the end of the sentence followed by a visual comprehension question in 50% of the trials. Participants answered the question by pressing "Yes" or "No" on the computer keyboard. Speed and accuracy were emphasized for the lexical decision and accuracy was emphasized for the comprehension question. There was an inter-trial interval of 500ms. Participants completed eleven practice trials with feedback before the actual experiment to familiarize with the procedure.

Hypotheses

All four groups were expected to show stronger priming effects for target words preceded by related sentence contexts and this effect was hypothesized to be stronger for native speakers (Table 26). The stress effect was expected to be stronger than the sentence context effect for Mandarin and Spanish speakers. In Spanish, the normal and most common stress pattern is for stress to fall on the final syllable of a word if it ends in a consonant and on the penultimate syllable if the final syllable ends in a vowel. Since over 70% of Spanish words end in a vowel, the predominant stress pattern in Spanish is penultimate, which becomes initial-stress in disyllabic words (Guion, Haraka & Clark, 2004). In addition, weak syllables cannot occur word-initially in Mandarin (Duanmu, 2007). Thus, it is hypothesized that both Spanish and Mandarin speakers would show stronger priming effects for SW words. The magnitude of priming effects between SW and WS words would not differ for Korean speakers due to the lack of word-stress contrasts in Korean. English speakers would show stronger priming effects for SW words since the predominant stress pattern in English disyllabic words is initial-stress (Cutler & Carter, 1987).

Table 26 Predictions (larger numeric magnitude represents greater weight given to that cue)

	English	Mandarin	Korean	Spanish
Sentence context	2	1	1	1
Stress cues	1	2	0	2

Data Analyses

Participants who scored lower than 60% accuracy for the comprehension questions were excluded from analyses. Only one Korean listener was excluded based on this criterion. In addition, five Korean speakers and one Spanish listener did not understand the requirement of the experiment and their accuracy for lexical judgments was lower than 60%. Data from these participants were also excluded. Therefore, data from 25 Korean, 21 Spanish, 30 Mandarin, and 35 English speakers were analyzed. The dependent variables were accuracy and (log-transformed) RT in the lexical decision task. RT data for incorrect responses were excluded (6.5% of total data). The independent variables were the stress location of the target words and the semantic relatedness between the sentence context and target. In the linear mixed-effects model predicting response latency and accuracy, priming condition, stress location, sentence context, and cloze test scores (as a covariate for proficiency) were the fixed effects while participant and item were the random effects (results from the model with language group as an additional fixed-effect is shown on Appendix H). Accuracy data was analyzed using the same mixed-effects model but with the binomial function. For brevity, only statistically significant results were reported.

Results

Table 27 shows the mean latencies and accuracy rates in the cross-modal priming lexical decision task. For brevity, only interactions between priming condition and the other factors (e.g., relatedness of sentence context and stress patterns) are reported since reliable interaction reflects the effect of sentence context or stress pattern on priming. In contrast to the findings in Experiment 1 in which speakers generally showed stronger priming effects to WS words, findings from Experiment 4 indicated greater priming effects for SW words (Figure 11). Cue interaction was observed in both Korean and Spanish speakers but not in English or Mandarin speakers. Korean and Spanish speakers showed greater priming effects for initial-stressed targets words than medial-stressed target words preceded by related sentence contexts; this stress effect was much smaller if the target words were preceded by unrelated contexts.

Table 27 Mean lexical decision RTs and % accuracy (in parentheses)

	English		Korean		Mandarin		Spanish	
Prime	SW	WS	SW	WS	SW	WS	SW	WS
Related context	701 (98.4)	727 (100)	987 (97.7)	1107 (98)	972 (99)	1156 (95)	903 (100)	1101 (93.5)
Baseline	792 (92.7)	795 (94.3)	1067 (89.9)	1105 (92)	1097 (99.1)	1174 (100)	1002 (97.6)	910 (93.5)
Mean priming	81	68	80	-2	125	18	99	-191
Unrelated context	708 (100)	735 (99.0)	1116 (100)	1029 (97.3)	987 (100)	995 (100)	833 (100)	1021 (98.5)
Baseline	784 (100)	789 (94.3)	1072 (92.3)	1045 (97.3)	1117 (95.7)	1081 (98.9)	916 (97.2)	1090 (97.1)

Mean priming	76	54	-44	16	130	86	83	69

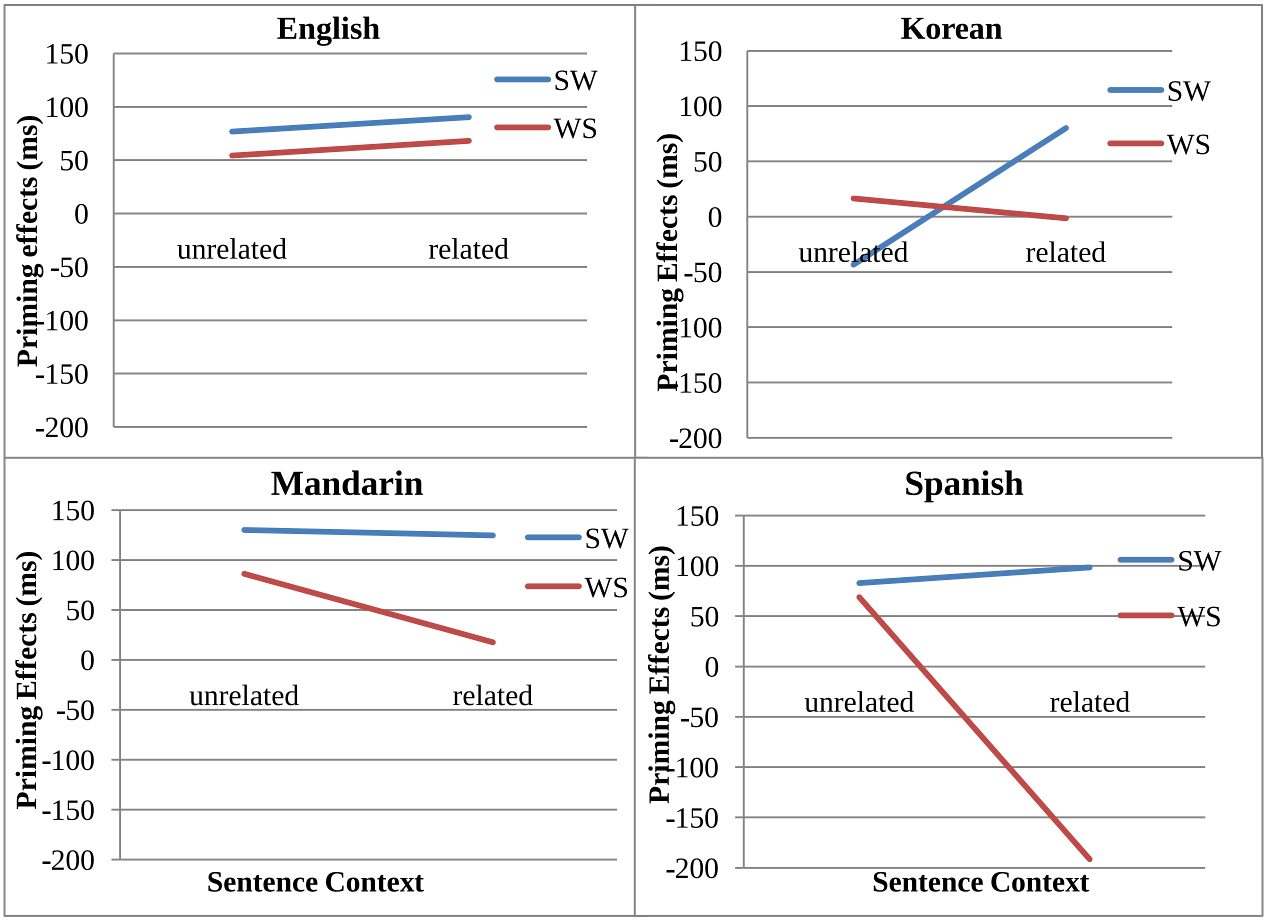

Figure 11 The use of stress location vs. sentence context (priming effects in RT)

English

Response Times. In the model predicting response latency (Table 28), the intercept estimated English speakers' log RT for initial-stressed target words preceded by unrelated sentence contexts. There was only a marginally significant effect of condition, suggesting that English speakers were marginally faster to make lexical decision for target words in the primed condition than those in the baseline condition.

Accuracy. In the model predicting accuracy, none of the effects or interactions reaches statistical significance.

Table 28 Mixed-effects linear model estimating log RT in the English group

Fixed Effects	Estimate	Std. Error	*t*-value	*p*MCMC
Intercept	3.099	0.210	14.742	0.0001
Condition (primed)	-0.035	0.019	-1.798	0.068
Sentence Context (related)	0.014	0.020	0.706	0.466
Stress (WS)	0.010	0.020	0.502	0.620
Cloze	-0.005	0.005	-1.124	0.141
Priming × Sentence context	-0.019	0.023	-0.82	0.422
Priming × Stress	0.002	0.024	0.078	0.918
Priming × Sentence context × Stress	0.010	0.032	0.309	0.769

Korean

Response Times. In the model predicting response latency (Table 29), there was a marginally significant three-way interaction among priming condition, sentence context, and stress location. As Figure 9 shows, for related sentence contexts, priming effects were larger when the target words were initial-stressed than when they were final-stressed; for unrelated sentence context, priming effects were greater for final-stressed target words than for initial-stressed target words.

Accuracy. In the model predicting accuracy, none of the effects or interactions reaches statistical significance.

Table 29 Mixed-effects linear model estimating log RT in the Korean group

Fixed Effects	Estimate	Std. Error	*t*-value	*p*MCMC
Intercept	3.083	.075	41.29	.0001
Condition (primed)	.001	.026	.05	.944
Sentence Context (related)	-.012	.038	-.4	.680
Stress (WS)	-.013	.039	-.35	.724
Cloze	-.002	.002	-.91	.305
Priming × Sentence context	-.040	.031	-1.31	.198
Priming × Stress	-.021	.031	-.67	.494
Priming × Sentence context × Stress	.076	.042	1.83	.076

Mandarin

Response Times. In the model predicting response latency (Table 30), there was a marginally significant effect of priming condition. Mandarin speakers were faster to identify target words in the primed conditions than those in the baseline conditions.

Table 30 Mixed-effects linear model estimating log RT in the Mandarin group

Fixed Effects	Estimate	Std. Error	*t*-value	*p*MCMC
Intercept	3.092	.117	26.453	.0001
Condition (primed)	-.042	.022	-1.887	.057
Sentence Context (related)	-.008	.036	-.23	.805
Stress (WS)	-.011	.037	-.288	.753
Cloze	-.002	.003	-.552	.548
Priming × Sentence context	-.005	.027	-.198	.849
Priming × Stress	.001	.027	.044	.952
Priming × Sentence context × Stress	.038	.036	1.045	.305

Accuracy. In the model predicting accuracy, there was only a significant effect of cloze test scores ($z = 2.11$, $p = .035$), suggesting higher L2 proficiency was associated with more accurate lexical judgments.

Spanish

Response Times. In the model predicting response latency (Table 31), there was a significant three-way interaction among priming condition, sentence context, and stress location. As Figure 9 shows, Spanish speakers showed stronger priming effects for initial-stressed target words than for final-stressed target words and this difference in priming effects was significantly larger in related sentence contexts than in unrelated sentence contexts. There was also a marginally significant effect of priming condition. Spanish speakers were faster to make lexical judgments for target words in the primed condition than those in the baseline condition.

Table 31 Mixed-effects linear model estimating log RT in the Spanish group

Fixed Effects	Estimate	Std. Error	*t*-value	*p*MCMC
Intercept	2.986	0.133	22.488	0.0001
Condition (primed)	-0.044	0.023	-1.87	0.069
Sentence Context (related)	0.033	0.030	1.085	0.265
Stress (WS)	0.067	0.031	2.143	0.035
Cloze	-0.001	0.003	-0.274	0.729
Priming × Sentence context	-0.005	0.028	-0.194	0.839
Priming × Stress	0.007	0.029	0.26	0.804
Priming × Sentence context × Stress	0.113	0.038	2.938	0.003

Accuracy. In the model predicting accuracy, there was no significant effect or interaction.

Comparison between Lexical Knowledge and Sentence Context

Post-hoc analyses were conducted to compare the relative importance of lexical knowledge and sentence context in L2 segmentation by combining the dataset from Experiments 1 and 4. In generalized linear mixed-effects models, condition (baseline or primed) cue type (two levels: lexical knowledge or sentence context), stress pattern (two levels: initial-stressed or medial/final-stressed), and cloze test scores were entered as fixed effects and participants and items were entered as random effects. For brevity, only significant interaction between priming condition and other factors are reported (Figure 12).

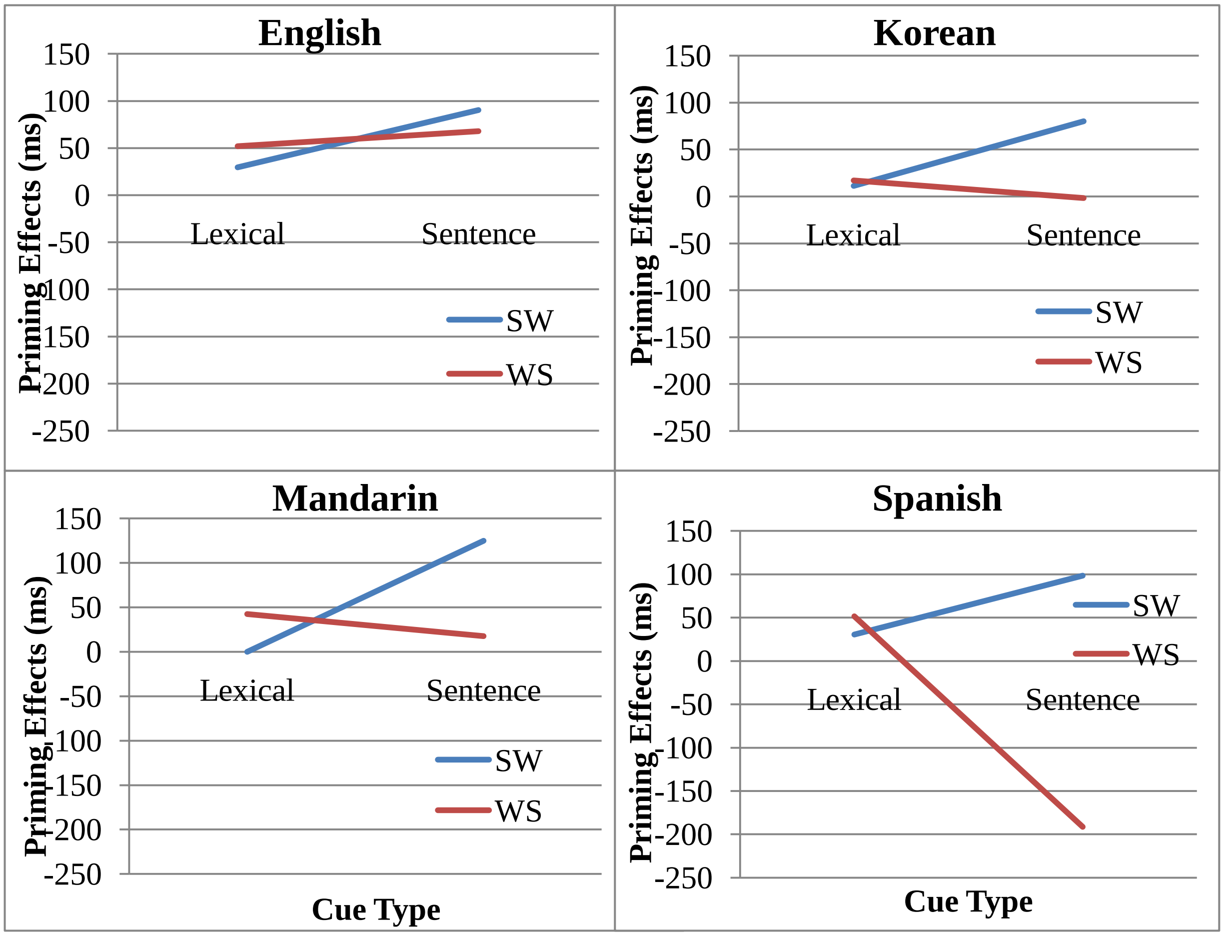

Figure 12 Comparison between lexical knowledge and sentence context (priming effects in RT)

English. The two-way interaction between condition and cue type was significant (t = -2.112, p = .033); priming effect was significantly stronger for sentence context cues than for lexical cues. The two-way interaction between condition and stress pattern was significant (t = -2.287, p = .027) when the cue type was lexical knowledge. As discussed in Experiment 1, English speakers showed stronger priming effects for medial-stressed words than for initial-stressed words in trisyllabic words. Although they showed the opposite pattern of result for the stress effect in disyllabic words Experiment 4, the three-way interaction among condition, cue type, and stress pattern was not significant (t = -2.112, p = .153).

***Korean*.** The two-way interaction between condition and cue type was significant (t = -2.04, p = .041). Similar to English speakers, Korean speakers showed stronger priming effects for sentence context cues than lexical cues. Priming condition did not significantly interact with any other variable (all ps > .1).

***Mandarin*.** Similar to both English and Korean speakers, Mandarin also showed a significant two-way interaction between condition and cue type (t = -2.593, p = .009). There was a marginally significant three-way interaction among priming condition, cue type, and stress pattern (t = 1.771, p = .079). Priming condition did not significantly interact with any other variable (all ps > .1). When the cue type was sentence context, priming effects were stronger for initial-stressed words than for final-stressed words; when cue type was lexical knowledge, the magnitude in priming effects was smaller between the two stress patterns.

***Spanish*.** There was a significant three-way interaction among priming condition, cue type, and stress pattern (t = 4.635, p = .0001). Similar to Mandarin speakers, Spanish speakers showed stronger priming effects for initial-stressed words for sentence context cues whereas the difference in magnitude of priming effects between the two stress patterns was significantly smaller for lexical cues. There were also two significant two-way interactions between priming condition and cue type (t = -2.243, p = .024) and between condition and stress pattern (t = -1.91, p = .056).

Discussion

Results of Experiment 4 are similar to those in Experiment 1. Both Korean and Spanish speakers showed interactive use of sentence context and stress pattern in L2 segmentation (though the effect was only marginally significant in the Korean group).

Both groups of L2 learners were significantly faster to segment initial-stressed words than final-stressed words in related sentence context; the difference in segmentation latency between the two stress patterns were smaller in unrelated sentence context. Both Mandarin and English speakers did not show any significant effect or interaction.

Stress Cues

None of the four language groups showed a significant interaction between priming condition and stress pattern, suggesting that priming effect did not differ significantly between the initial-stressed and medial-stressed words. Thus, Experiment 4 did not find any empirical support for the independent use of stress cues in all four language groups (see below for discussion of cue interaction).

Sentence Context

There was no significant interaction between priming condition and sentence context in all four language groups, suggesting that related sentence context did not result in faster identification of the target words compared to unrelated sentence context. Thus, Experiment 4 did not find any empirical support for the independent use of sentence context in any of the language groups.

Cue Interaction

Similar to the results in Experiment 1, English speakers did not show any trend of cue interaction. As Figure 10 shows, the two lines are perfectly parallel for the English group. In contrast, all three L2 group show a trend of cue interaction, although the interaction was only significant in the Spanish group. Since L2 learners did not behave similarly to native English speakers, it is likely that L1 prosodic characteristics influence the interaction between stress cues and sentence context. When sentence context has high predictability for the occurrence of the target word, Spanish speakers showed faster segmentation for initial-stressed words than final-stressed words; when sentence context has low predictability for the occurrence of the target word, the difference in segmentation latency was smaller between the two stress patterns. It is possible that listeners were building up and revising their predictions for the upcoming target word as the speech signal unfolds overtime and in the related sentence context condition, the occurrence of the target word would have a high probability of being consistent with their predication and thus facilitates the identification of word boundaries. This process may be further accelerated when participants hear initial-stressed target words in which phonetic prominence coincides with word edges (for Korean speakers) and the stress pattern is consistent with the predominant pattern in L1 (for Mandarin and Spanish speakers). In fact, the condition with initial-stressed target words and related sentence context produced the strongest priming effects for Korean and Spanish speakers and the second strongest priming effects for Mandarin speakers.

When participants heard a final-stressed word in which stress location does not coincide with the beginning of the word, they may initially consider the unstressed initial-syllable in the target word as the final-syllable of the preceding word. For example, in the auditory sentence fragment "Her parents are religious Christians, they taught her to believe," participants may initially parse the target word as "to be leave." When participants realized they had placed the word boundary in the wrong location, it would take them longer to make the revision because it would be difficult to overcome the previous predication as L2 learners have been shown to be less reluctant to revise their initial interpretation (Field, 2008). In contrast, when the sentence context was less predictable for the occurrence of the target words, it may be easier for participants to overcome any prediction that was inconsistent with the speech input since their predication might have been built with a slower pace or weaker strength. Thus, in the unrelated sentence context condition, the difference in segmentation latency between initial- and final-stressed words was not as large as that in the related sentence context condition.

Cue Type: Lexical Knowledge vs. Sentence Context (Experiments 1 and 4)

All four language groups showed stronger priming effects for related sentence context than for real word context, suggesting that both native and nonnative speakers weigh sentence context over lexical knowledge in speech segmentation. The difference in cue weighing is particularly prominent in initial-stressed words (Figure 11), though this interaction between cue type and stress pattern was only significant in the Spanish group. Words with initial-stress may be easier to segment because it coincides with beginning of the word and it is consistent with the predominant stress pattern in Mandarin and English words and in Spanish disyllabic words. The relative importance of sentence context compared to lexical knowledge is consistent with the Hierarchical Framework (Mattys et al., 2005) which postulates that semantic and syntactic content of the utterance contributes to segmentation and subsequent lexical access by favoring those words most likely given a particular context. Since it is unlikely that the short list of word candidates would include nonwords, the knowledge of what constitutes a real word in English is not as important if participants can rely on sentence context to identify word boundaries.

Chapter 8 General Discussion

This dissertation project examined the use of segmentation cues by L2 learners of English with Korean, Mandarin, or Spanish L1s. Four language groups of participants took part in four experiments which tested the predications of the Revised Framework for L2 segmentation (RFL2). The RFL2 predicts that L2 learners would rely more on sublexical cues (e.g., word stress and acoustic-phonetic cues) than lexical cues (e.g., lexical knowledge and sentence context) to segment L2 speech. Experiment 1 compared the use of stress cues and lexical knowledge in a cross-modal primed lexical decision task. Results showed that native English speakers used stress cues and lexical knowledge independently whereas Korean and Spanish speakers used the cues interactively. When lexical cues are available (i.e. the target word for segmentation is preceded by a real word context), Spanish speakers, but not Korean speakers, appeared to use stress cues. When lexical cues are absent (i.e. the target word is preceded by a nonword context), Korean speakers seemed to use stress cues but not Spanish speakers. However, these results did not indicate whether one cue is given more weight than the other.

Experiment 2 compared the use of phonotactic cues and lexical knowledge in a word spotting task. English speakers showed evidence of using lexical knowledge only in the coda /ŋ/ condition. English, Korean, and Mandarin speakers showed evidence of using acoustic-phonetic cues. Coda /s/ makes the identification of word boundaries more difficult probably because it has weaker acoustic qualities in English and the nonnative listeners have developed sensitivity to these allophonic variations in L2. Lexical knowledge may have helped compensate for some of the difficulties with segmenting words with coda /s/ since the coda /s/ effect is only significant in nonword context condition for the L2 learners, The overall result from Experiment 2 suggest that acoustic-phonetic cues are more important than lexical knowledge for Mandarin and Korean speakers. There was no indication that native English speakers weight one cue over the other.

Experiment 3 compared the use of phonotactic cues and word-level semantic cues in a word-spotting task. English, Mandarin, and Spanish speakers all showed a coda /s/ effect similar to that found in Experiment 2. English speakers also showed evidence of using semantic cues in the coda /n/ condition while Korean speakers showed cue interaction. While English and Mandarin speakers' difficulty with segmenting words with coda /s/ may stem from the less robust acoustic qualities of fricatives compared to nasals, Spanish speakers were also influenced by the phonological variation of word-final /s/ in their L1. This result indicates the influence of L1 acoustic-phonetics on L2 segmentation for Spanish speakers. Since no significant effect of semantic relatedness was observed in Mandarin and Spanish speakers, it appeared to both L2 groups rely more on acoustic-phonetics than word-level semantics. Comparison of the data from Experiments 2 and 3 suggest that all three L2 groups, but not native English speakers, weight word-level semantics over lexical knowledge.

Finally, Experiment 4 compared the use of stress cues and sentence context in a cross-modal primed lexical decision task. Spanish and Korean speakers showed trends of cue interaction. Stronger priming effects were observed for initial-stressed words and this stress effect was greater when the target word was preceded by related sentence context than when it was preceded by unrelated sentence context. Comparison of the data from Experiments 1 and 4 suggest that all four language groups weight sentence context over lexical knowledge. In Spanish, the predominant stress location is penultimate for trisyllabic words and initial for disyllabic words. Stronger priming effects for medial-stressed words in Experiment 1 and stronger priming effects for initial-stressed words in Experiment 4 indicates the influence of L1 stress characteristics on L2 segmentation.

Revisiting the Revised Framework for L2 Segmentation

The RFL2 makes five predictions: 1) learners with low to intermediate L2 proficiency will weigh sublexical cues over lexical cues; 2) among sublexical cues, segmental cues will be relatively more important than prosodic cues for L2 learners; 3) learners whose L1 does not have word-level stress will not be able to utilize stress cues to segment an L2 with word stress; 4) among lexical cues, learners will weigh lexical knowledge over sentence context; 5) within the segmental level, acoustic-phonetic cues outweigh phonotactic cues. The hypothesized cue weighing for the four types of cues tested in this dissertation is shown in Table 32. The proposed weighting for native English speakers is based on the Hierarchical Framework (Mattys et al., 2005) while the weighting for the L2 groups is based on the RFL2.

Table 32 Hypothesized cue weighting for the four language groups (larger number indicates greater weight)

Cue Type	English	Korean	Mandarin	Spanish
Sentence Context	4	1	1	1
Word-level semantics	4	1	1	1
Lexical Knowledge	3	2	2	2
Phonotactics/Acoustic-Phonetics	2	3	4	4
Word Stress	1	0	3	3

Revised Cue Weightings

The first hypothesis of RFL2 regarding the relative weights of sublexical and lexical cues was tested in all four experiments. In Experiments 1 and 4, cue interaction was observed in both Korean and Spanish groups, suggesting that stress cues were not given more weight than lexical knowledge or sentence context. In Experiment 2, both Mandarin and Korean speakers appeared to weigh segmental cues (acoustic-phonetics or probabilistic phonotactics) over lexical knowledge. In Experiment 3, both Mandarin and Spanish speakers seemed to rely more on segmental cues than word-level semantics while Korean speakers showed cue interaction. The third hypothesis regarding Korean speakers' use of stress cues was tested in Experiments 1 and 4. Korean speakers appeared to be able to utilize pitch prominence in stressed syllables to segment L2 speech even though they do not encode stress in lexical access as shown in a previous study (Lin et al., under review). The fourth hypothesis regarding the relative weights of sentence context and lexical knowledge was tested by comparing the data in Experiments 1 and 4. All four language groups appeared to weigh sentence context over lexical knowledge. In addition, comparing the data in Experiments 2 and 3 also suggests that all three L2 groups relied more on word-level semantic cues than lexical knowledge. The second and fifth hypotheses were not directly tested in any experiment. Table 33 shows the revised cue weighting based on the results of the four experiments.

Table 33 Revised cue weightings for the four language groups based on results of Experiments 1-4

	Lexical Cues						Sublexical Cues	
English	Sentence Context	>	Word-level Semantics	=	Lexical knowledge	< ?	Acoustic-Phonetics/ Phonotactics	Word Stress √
Korean	Sentence Context	>	Word-level Semantics	>	Lexical knowledge	<	Acoustic-Phonetics/ Phonotactics	Word Stress √
Mandarin	Sentence Context	>	Word-level Semantics	>	Lexical knowledge	<	Acoustic-Phonetics/ Phonotactics	Word Stress ?
Spanish	Sentence Context	>	Word-level Semantics	>	Lexical knowledge	<	Acoustic-Phonetics/ Phonotactics	Word Stress √

Note. ? indicates uncertainty in the weight or use of cues due to inconclusive findings

For native English speakers, the three clear results were that 1) sentence context weighs over lexical knowledge; 2) word-level semantic cues do not weigh over lexical knowledge; and 3) stress cues were used in native segmentation independently. In Experiment 2, English speakers showed evidence of using lexical knowledge only in the coda /ŋ/ condition but not in the /n/ or /s/ condition. English speakers may only use lexical knowledge if a coda cue has a 0% phonotactic probability at word-initial or word-final positions. Based on this reasoning, English speakers may weigh probabilistic phonotactics over lexical knowledge. However, this conclusion is only tentative since in the current study phonotactic cues were operationalized as absolute legality (i.e. whether a phoneme is allowed in a certain position in a syllable) rather than relative probability (i.e. how likely a phoneme is to occur in a certain position in a syllable). In both Experiments 1 and 4, there was no empirical evidence indicating that stress cues are relatively less important than lexical knowledge or sentence context. The only safe conclusion to be drawn is that English speakers do use stress cues even when lexical cues are available.

For Korean speakers, the clear results were that 1) they are capable of using stress cues (more precisely, phonetic prominence); 2) segmental cues weigh over lexical knowledge; and 3) sentence context is relatively more important than word-level semantics, which in turn is relatively more important than lexical knowledge. Based on these results, the segmentation cue with the strongest weight in the lexical tier is postulated to be sentence context, followed by word-level semantics and lexical knowledge. Segmental cues are relatively more important than lexical knowledge. Korean speakers appeared to be able to utilize the pitch cues in stressed syllables to segment L2 speech. Although not shown in Table 33, Korean speakers used cues at the lexical level and those at the sublexical level interactively (see Figure 13).

For Mandarin speakers, the clear results were that 1) segmental cues weigh over lexical knowledge and semantic cues; and 2) sentence context is relatively more important than word-level semantics, which in turn are more important than lexical knowledge. Based on these findings, sentence context was considered as the strongest cue in the lexical level, followed by word-level semantics and lexical knowledge. In Experiments 1 and 4, Mandarin speakers did not show any significant stress effect or interaction involving stress location. Although Mandarin has lexically contrastive stress (Duanmu, 2007) and previous research has shown that native Mandarin speakers encode stress in the processing of L2 English words (Lin et al., under review), it is not clear in the current study whether Mandarin speakers utilize stress cues to segment L2 English speech.

Spanish speakers showed a similar pattern of cue weighting at the lexical level like Korean and Mandarin speakers did. In addition, Spanish speakers appeared to weigh segmental cues over word-level semantics. Spanish speakers also showed the influence from L1 phonological variation of word-final /s/ on their use of this coda cue in L2 segmentation. Their use of stress cues interact with lexical knowledge and sentence context. Thus, no clear conclusion can be drawn regarding the relative weight of stress cues compared to the cues at the lexical level. The only safe conclusion to be drawn is that Spanish speakers do use stress cues even when lexical cues are available.

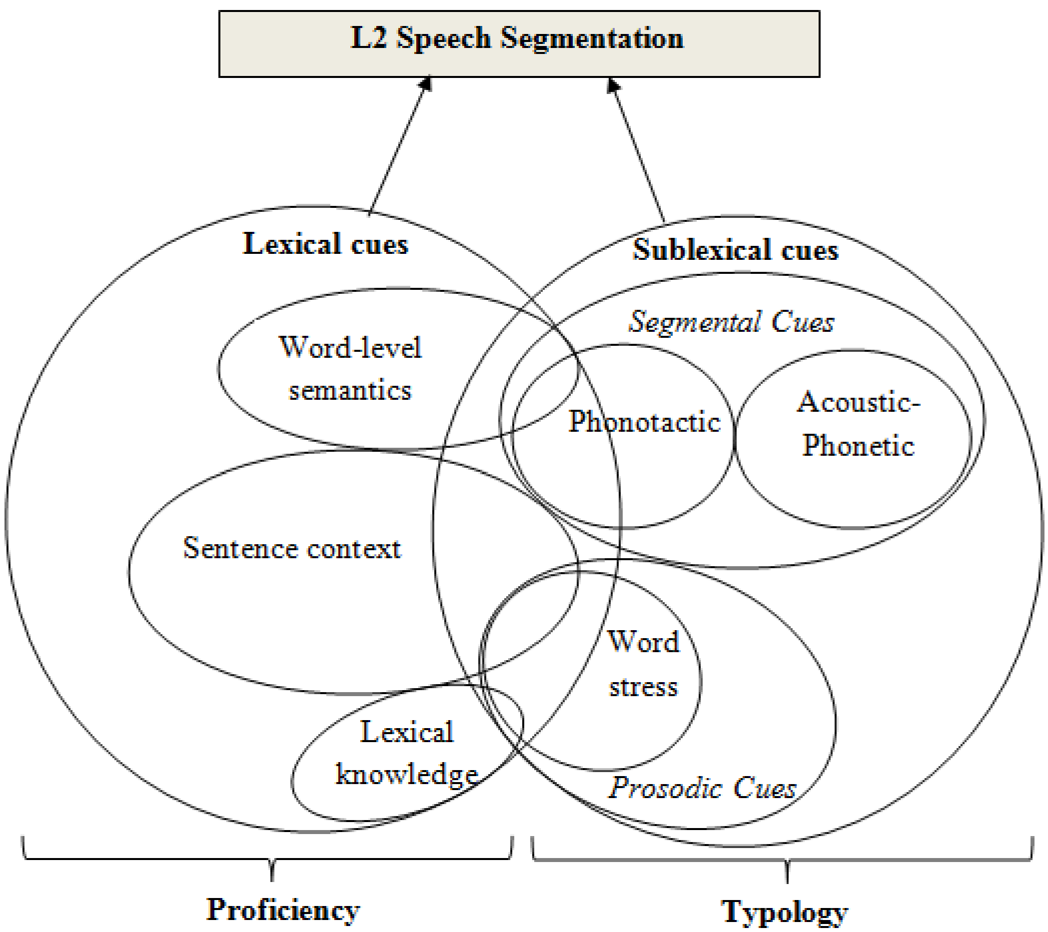

Figure 13 Revised Framework for L2 Segmentation based on results

In light of the revised cue weightings, small modifications were made to RFL2 to reflect the results of the current dissertation (Figure 13). The area of the circle represents the relative importance of that particular cue in L2 speech segmentation. The updated version of RFL2 predicts that L2 learners would give similar weights to lexical and sublexical cues. More importantly, the overlap of the circles indicates that there will be interaction between the two levels of cues. Particularly, stress cues would interact with sentence context and lexical knowledge (i.e. Korean and Spanish speakers) and phonotactic cues would interact with semantic cues (i.e. Korean speakers). Within the lexical cues, RFL2 predicts that sentence context would be relatively more important than word-level semantics, hence the larger area of the circle for sentence context. In addition, word-level semantics would be more important than lexical knowledge in L2 segmentation. The updated RFL2 considers word-level semantics as an individual subset of lexical cues rather than grouping them under a broad category of sentence context since semantic relatedness at the word level is weighed differently compared to sentence context and lexical knowledge by L2 learners. With regards to the prosodic cues, it appeared that even L2 learners whose L1 does not have contrastive stress can use phonetic distinction in stressed syllables to segment L2 speech (i.e. Korean speakers). This is indicated by changing the circle of word stress from a dashed line to a solid line. Finally, RFL2 predicts that segmental cues would be more important than lexical knowledge (i.e. Mandarin and Korean speakers in Experiment 2).

Cross-Linguistic Influences on L2 Segmentation

RFL2 postulates that cue weighing at the sublexical level would vary by the characteristics of the phonological system in the learner's L1. For the use of acoustic-phonetic cues, for example, the phonological variation of word-final /s/ in Spanish contribute to Spanish speakers' difficulty with segmenting words with coda /s/ than those with coda /n/ and this L1 influence is additive to the difficulty associated with less robust acoustic qualities of fricatives. This finding is consistent with previous research (Weber, 2000) which found that advanced German L2 learners of English experienced more difficulty segmenting L2 English words beginning with consonant clusters that are phonotactically illegal onsets in German. In addition to position-specific phonotactic constraints, previous research has also shown that possible-word constraints (i.e. the result of segmentation must be a possible word in the listener's language, Norris et al., 1997) in L1 influence segmentation in L2. Slovak, a West Slavic language, allows words consisting only of a single consonant whereas German does not. Native Slovak L2 learners of German found it harder to recognize the target German word *rose* in the sequence *trose* than in *krose* since *k*, not *t*, is a real word in Slovak (Hanulikova, Mitterer, & McQueen, 2011). Altenberg (2005) showed that native Spanish speakers were better at using allophonic cues that exist in their L1 (i.e. glottalization in word-initial vowels) to segment L2 English speech than using allophonic cues that do not exist in their L1 (i.e. aspiration in word-initial stops). Previous research along with the current study has consistently demonstrated that allophonic variations or phonotactic constraints in L1 play an important role in the use of segmental cues in L2 segmentation.

On the other hand, there is also evidence showing that L2-specific phonotactic constraints can be learned by L2 learners and used in nonnative segmentation. Weber and Cutler (2006) found that highly proficient German L2 learners of English identified the target word *lecture* faster in the nonword sequence *thrarshlecture* than in the sequence *gorkleccture* where both /ʃl/ and /kl/ are legal in German whereas /ʃl/ is not a legal onset in English. Similar results have been found in native Arabic L2 learners of English (Al-jasser, 2006). In Experiments 2 and 3 of the current dissertation, Spanish speakers did not appear to experience any difficulty using coda /ŋ/ to segment L2 English speech even though this phoneme does not exist in L1. The current findings along with previous research suggest that while L1 knowledge remains active in nonnative listening, L2 learners may be able to acquire L2 phonotactics and use them efficiently in L2 segmentation.

For the use of stress cues, Spanish speakers were faster to segment words with a stress pattern that is consistent with the predominant pattern in their L1. Even though Korean speakers' L1 does not have lexically contrastive stress, Korean speakers were able to segment L2 words by using phonetic distinctions such as final F0 rise and final lengthening in stressed syllables as a result of experiences with the prosodic phrasal pattern in L1. This finding is consistent with previous research (i.e. Sanders et al., 2002) which has shown that native Japanese speakers were able to use stress cues to segment L2 English speech. Japanese is a pitch accent language and does not have lexically contrastive stress. Sanders et al. (2002) found that Japanese speakers were faster to identify a target phoneme in a stressed syllable than the same phoneme in an unstressed syllable.

It should be noted that Spanish and Korean speakers showed different pattern of use of stress cues. In Experiment 1, Spanish speakers were faster to segment medial-stressed words compared to initial-stressed words in *real word* context whereas Korean speakers were faster to segment medial-stressed words in *nonword* context. Spanish have lexically contrastive stress and Spanish speakers encode stress in lexical access in L1 (Dupoux et al., 2001). While Spanish speakers' sensitivity to medial-stressed words may be influenced by the predominant stress pattern in Spanish (i.e. penultimate), Korean speakers' use of stress cues is possibly affected by the typical pitch pattern of the accentual phrase (i.e. Low High Low High) in Korean. Although speakers whose L1 does not have contrastive stress may be able to use phonetic prominence to segment L2 speech, the mechanism underlying their use of stress cues is likely to be different from that of speakers whose L1 does have contrastive stress.

Findings of the Korean speakers in the current dissertation with regards to the use of stress cues are not consistent with a previous study involving native Hungarian L2 learners of English (White et al., 2010). Using a cross-modal primed lexical decision task similar to Experiment 1 of the current dissertation, White et al. did not observe a significant interaction between stress pattern and lexicality in priming effects. In addition, there was no significant interaction between priming condition and stress pattern, suggesting that native Hungarian speakers did not utilize stress cues to segment L2 English speech. Hungarian is a fixed-stressed language in which stress is always word-initial (Siptar & Torkenczy, 2000). If Hungarian speakers applied L1 stress patterns to L2 segmentation, identification of word boundaries should be easier for initial-stressed words than for medial-stressed words. Similar to Hungarian, Slovak also has fixed-stress in word-initial position. Hanulikova et al. (2011) found that native Slovak speakers did not apply L1 stress pattern to segment L2 German, a language with contrastive stress like English. These results suggest that Slovak speakers have suppressed native segmentation strategies and adapt to a particular segmentation procedure appropriate for German input.

Considering the inconsistency of the findings with regards to L2 learners' use of stress cues, it appears that L1 phonotactic and possible-word constraints may have a stronger cross-linguistic influence on L2 segmentation than L1 stress patterns.

Compared to the use of sublexical cues, the use of lexical cues does not seem to rely on typology. Despite the vast differences in the phonological systems of Korean, Mandarin, and Spanish, all three groups of L2 learners showed a consistent pattern of cue weighting in which sentence context was given the strongest weight, then followed by word-level semantics and lexical knowledge. L2 learners' weighting of the lexical cues is similar to that of native English speakers who weighed sentence context over lexical knowledge. The L2 learners in the current study may be considered as intermediate to advanced learners considering they started English acquisition at an early age of 10 (on average) and they scored 33 or above out of 50 in the cloze test (English speakers scored 46 out of 50). It is possible that as L2 learners' English proficiency improved, their use of lexical cues in L2 segmentation becomes more native-like. However, we must be cautious with these interpretations given that the different stimuli and design were used across experiments.

One critical difference between native and nonnative listeners' use of segmentation cues is that native English speakers appear to use lexical and sublexical cues independently while L2 learners tend to show cue interaction. L2 learners tend to be slower in lexical access than native speakers. Although they could utilize lexical knowledge to identify word boundaries, they are slower to settle on a lexical entry and the information about stress location in real words bootstrap this word identification process. Research in the L2 acquisition literature has demonstrated that L1 phonological information can bootstrap L2 learners' word recognition during reading (see Koda, 2004 for a review). It is very likely that phonological information (i.e. word stress or acoustic-phonetics) in L1, L2, or a combination of both, can bootstrap word segmentation in L2 speech comprehension.

Native English Segmentation

The results of the native English speakers in the current study are not consistent with those in Mattys et al. (2005). In the comparison between the use of stress cues and lexical knowledge in Experiment 1, the current study showed that native English speakers used both cues independently and there was no evidence indicating one cue was weighed over the other. In contrast, Mattys et al. (Exp 3) found that native English speakers only used lexical knowledge in optimal listening condition; when the auditory stimuli was masked by noise and lexical information was not available, English speakers then use stress cues. These results motivated Mattys et al. to propose the Hierarchical Framework which postulates that segmentation is lexically driven. However, the current study found that both lexical and sublexical information contributes to L1 English segmentation.

There are a number of possible reasons for these inconsistent findings. First, different stimuli were used. In the current study, mostly high frequency words were chosen so that the L2 learners would not consider them as nonwords. Mattys et al. (Exp 3) included some low frequency words such as *vaccinate* and *versatile*. It is possible that lexical knowledge may become more important when listeners encounter speech input with a variety of frequency. Second, the participants in Mattys et al. are speakers of British English while the participants in the current study are speakers of American English. Cutler and Carter (1987) suggested that American English has a tendency to favor initial stress more than British English does. As a result, there are more initial-stressed polysyllabic words in American English than in British English. Thus, stress location may be a relatively more reliable cue to word boundary in American English. Third, the current study did not examine the use of cues under poor listening condition. In Experiment 1 of the current dissertation, it was found that English speakers use both stress location and lexical knowledge independently under optimal listening condition. This pattern of cue use may change when noise is added to the auditory stimuli.

Despite these differences, findings in the current study bring into question the generalizability of the Hierarchical Framework. It is possible that sublexical cues were not only used when lexical information is unavailable, impoverished, or ambiguous. When all cues are equally available in the speech signal, English speakers may consider both lexical and sublexical cues as long as they are not providing conflicting information about word boundary.

When considering other existing word segmentation theories in the field, native English speakers' use of stress cues in the current study only partially supported the Metrical Segmentation Strategy (MSS, Culter & Norris, 1988) hypothesis which postulates that English speakers segment speech at the onset of every strong syllable. In the context of Experiment 1 in the current dissertation, MSS predicts that English speakers would show stronger priming effects for initial-stressed words than medial-stressed words. However, stronger priming effects were observed for trisyllabic medial-stressed words. Cutler and Carter (1987) observed that 70% of English disyllabic content words begin with a strong syllable. However, the percentage of initial-stressed words is smaller in trisyllabic words. In a database of 20,000 English words, Clopper (2002) identified a total of 4498 trisyllabic words, in which 2619 of them have initial stress, 1510 have medial stress, and 369 have final stress. The proportion of initial-stressed trisyllabic words in this corpus is 2619 ÷ 4498 = 58%, as opposed to 70% initial-stressed disyllabic words in Cutler & Carter (1987). Native English speakers may not use MSS consistently for trisyllabic words since segmenting at the onset of every stressed-syllable would result in larger number of errors compared to using MSS to segment disyllabic words.

Limitations

The RFL2 operates under the premises that all cues are simultaneously available in the speech signal and speakers have the ability to utilize all available cues in the signal. However, both of these premises may not be always true in the natural speech environment for L2 learners. First, there are some cues that may become available earlier than others. For example, segmental cues at word-initial positions arrive earlier than those at word-final positions due to the sequential nature of spoken speech. Word-level cues are often available earlier than sentence-level cues as speakers can access the semantic representations of individual words but cannot integrate their meanings until more words are revealed in the spoken input. Furthermore, early-arriving cues may have an inhibitory effect on late-arriving cues as L2 learners have been shown to be less reluctant to revise their initial interpretation (Field, 2008). Previous research has shown that L2 learners, compared to native speakers, are slower to recover and revise their hypotheses when processing garden path sentences (e.g., *The horse raced pass the barn fell*) (Roberts & Felser, 2011). Thus, L2 learners may also be less inclined to use sentence-level cues in speech segmentation if it requires ambiguity resolution. However, the current dissertation is not designed to examine the influence of the timing of cue availability on L2 segmentation.

L2 learners' difficulty with on-line recovery from temporarily ambiguous sentences may be due to their relying on lexical-semantic, but not syntactic, information in sentence processing. Since semantic and syntactic information is utilized differently by L2 learners, a more fine-grained model for L2 segmentation may further divide the sentence context cues into two sublevels: word semantics and sentence syntax. According to the Shallow Structure Hypothesis proposed by Clahsen and Felser (2006), the representations that L2 learners compute for sentence comprehension are shallower and less detailed than those computed by native speakers. Nonnative speakers tend to rely on lexical-semantic parsing and each new incoming chunk of information is integrated in the emerging semantic representation incrementally. L2 learners did not show any evidence of structure-based parsing like native speakers (Felser, Roberts, Marinis, & Gross, 2003). For example, for the sentence *The dean liked the secretary of the professor who was reading a letter*, native English speakers showed a preference of linking the preposition to the second noun phrase (NP2) whereas L2 learners with Greek or German L1 did not show a preference for either NP1 or NP2. These results suggested that L2 learners may be less sensitive to syntactic structures in online sentence processing. In speech segmentation, it is possible that L2 semantic cues may be more accessible and weighed over L2 syntactic cues. However, the use of syntactic cues was not examined in the current dissertation and its relative importance compared to other lexical cues in L2 segmentation warrants further studies.

In addition to the differentiation between L2 semantic and syntactic cues, the current dissertation also did not compare L1 segmental cues and stress cues. This is because many segmental cues are confounded with stress. For example, allophonic cues such as increased aspiration in word-initial voiceless stops (Christie, 1974; Nakatani & Dukes, 1977) are useful segmental cues for segmentation. However, stressed-syllables usually have longer VOT than unstressed syllables, which would result in the perception of stronger aspiration in word-initial voiceless stops in stressed syllables. Similarly, word-initial stressed vowels are glottalized more frequently than word-initial unstressed vowels (Pierrehumbert, & Frisch, 1997). Mattys (2004, et al., 2005) have used coarticulation to examine the use of acoustic cues in segmentation. However, de Jong, Beckman, and Edwards (1993) found that coarticulation is reduced in stressed syllables because stressed syllables tend to be hyperarticulated. This confound between coarticulation and stress cues casts doubt on Mattys' conclusion that acoustic cues are more important than stress cues in native English segmentation. Due to the inability to identify an acoustic cue that is independent of stress, the current dissertation was not able to the second hypothesis of RFL2 which predicts that L2 learners assign more weight to segmental cues than to word stress.

Due to time constraints, many fine-grained cross-linguistic differences were not examined in the current dissertation. In Experiments 2 and 3, three codas, /n ŋ s/, were selected to examine the use of phonotactic cues in L2 segmentation. Both Mandarin and Korean have codas /n ŋ/ and it was predicted that L2 learners from these two language groups will behave similarly in terms of utilizing the coda cue to identify word boundaries. Unlike Mandarin which only allows two phonemes in the coda position, Korean has a slightly larger coda inventory allowing /p, t, k, m, l/ at word-final positions in addition to /n, ŋ/. Due to its more restrictive phonotactic constraint for codas, Mandarin speakers may rely more on coda cues compared to Korean speakers. Thus, the weight assigned to L1 phonotactic cues by Mandarin speakers may be different from that by Korean speakers.

Among the three L1s examined in this dissertation, Mandarin is considered to be typologically closest to English in terms of word stress. However, there exist many fine-grained differences between the English and Mandarin stress systems. In Mandarin, a weak syllable cannot be word-initial (Duanmu, 2007) whereas in English a weak-syllable can occur in any location in a word (e.g., word-initial: *support* /səˈpɔrt/, word-medial: *catalog* /ˈkætəˌlɒg/, and word-final: *museum* /mjuːˈzɪəm/). One of the acoustic correlates of stress realization is pitch and higher pitch tends to lead to more identification of stress by native English speakers (Lin, Lukyanchenko, Winn, & Idsardi, in preparation). However, Shen (1993) found that Mandarin speakers often rely more on duration and intensity rather than pitch in stress perception because pitch contour constitutes one of the dimensions of tonal identification. The current dissertation is mainly concerned with the use of stress location as a cue for word boundaries. It remains to be seen whether the fine-grained differences in the acoustic correlates of stress realization may or may not have an influence on the use of stress cues.

Finally, there are a few limitations associated with the sample. Mattys et al. (2005) tested an average of 50 participants in each of the experiments whereas the current study only included an average of 30 participants in each language group. The lack of significant effects in some of the experiments may be due to the small sample size.

Participants in the Mandarin and Korean groups were recruited and tested in Los Angeles whereas participants in the English and Spanish groups were recruited and tested in College Park. The Spanish group had higher proficiency (as measured by the cloze test) and larger amount of L2 exposure than both the Mandarin and Korean groups. Since there is a larger population of native Mandarin and Korean speakers in Los Angeles CA than in College Park MD, the Mandarin and Korean speakers may use English less frequently than the Spanish speakers. However, the Mandarin group's result pattern is more similar to that of the Spanish group rather than the Korean group, suggesting that the location of testing and amount of L2 exposure have minimal, if any, impact on the use of L2 segmentation cues. The Mandarin and Spanish groups are more heterogeneous in terms of L1 background than the Korean group. The Spanish speakers came from 10 different Spanish-speaking countries while the Mandarin speakers came from 8 different Northern provinces or cities in Mainland China. There could be some fine-grained differences in the particular dialect of Mandarin or Spanish they speak. However, it is unlikely dialectal differences would influence the use of those segmentation cues examined in the current study since the phonotactic constraints and stress characteristics are still applicable regardless of which dialect of Mandarin or Spanish the participants speak.

Direction for Future Research

Unlike Mattys et al. (2005), the current dissertation did not consider the role of listening condition in L2 segmentation. The auditory stimuli presented in all four experiments were intact and unambiguous. L2 learners' weighting of cues may change if noise is added. Since the current study found interactive use of lexical and sublexical cues, an interesting research question to investigate is, when lexical information is masked by noise, whether L2 learners would still be able to independently rely on sublexical cues (particularly, stress location) to segment continuous speech in L2.

It has been found that acoustic cues outweigh absolute phonotactics in native English segmentation (Newman et al., 2011). Results from Experiment 2 and 3 of the current dissertation showed that English speakers were slower to segment target words with coda /s/ than those with coda /n/. Although both phonemes are legal in word-final position, /s/ has weaker allophonic variations and this may make it more difficult to identify word boundaries following /s/. However, /s/ has similar relative frequency in word-initial and word-final positions whereas /n/ has a higher likelihood of occurrence in word-final than in word-initial position. L2 speakers also showed a similar pattern of results. This may be due to the weaker allophonic strength of /s/ or the different phonotactic probabilities of word-final /n/ and /s/ in English. The current study cannot tease apart these possibilities. In future studies, researchers should choose a pair of onsets or codas that have similar phonotactic probability in L1 but vary in the strength of allophonic variation in L2 to examine whether the results found in native English speakers can be replicated in L2 segmentation.

Another direction for future research is to compare the relative weighing of cues such as syntax and pragmatics at the sentence level. In Experiment 4 of the current dissertation, sentence context was operationalized as the likelihood of occurrence of the target word based on the semantic context of the preceding sentence fragment. Manipulating the syntactic structure of the preceding sentence fragment may also influence the likelihood of occurrence of the target word. Based on Clahsen and Felser's

(2006) Shallow Structure Hypothesis, L2 learners may rely more on the meaning of the sentence rather than its syntactic structure to segment L2 speech. If this is true, participants should segment the target words that are consistent with the meaning of the sentence context faster compared to those that are not consistent, regardless of the syntactic complexity of the sentence structure.

Finally, more L2 learners with different L1 backgrounds should be tested to examine the generalizability of RFL2. It is very likely that the relative weighting of the cues at the sublexical level varies depending on learners' L1. For example, Thai and Vietnamese are both tonal languages without lexically contrastive stress (Wayland, Guion, Landfair, & Li, 2006; Nguyen& Ingram, 2005). English learners with these L1 backgrounds may be able to rely on pitch cues in stressed syllables to segment L2 speech like the Korean speakers in the current study did. Similar to Spanish, Japanese does not have /ŋ/ (it is an allophone of /n/) in its phoneme inventory. Perhaps Japanese speakers who learn English as L2 would be able to use this coda in L2 segmentation like the Spanish speakers in the current study did.

Broader Impact

This dissertation project has significant implications for pedagogical practices. Understanding speech is a critical component in language acquisition and identifying cues that can facilitate this process will benefit both learners and teachers. Findings from this dissertation may help inform teachers which cues are employed more reliably by L2 learners from various L1 backgrounds. L2 speech comprehension will be more effective if teachers can selectively direct students' attention to cues in the L2 that are not utilized in L1 segmentation (e.g., the use of initial-stress to signal word onset for Spanish and

Korean speakers). It has been shown that differences between learners' L1 and L2 create difficulties in the learning process. Japanese speakers have difficulty discriminating between the English words *light* and *right* because the /r-l/ distinction does not exist in Japanese (Iverson, et al., 2003). Japanese listeners' perception of English /r-l/ can be improved through phonetic training (Iverson, Hazan, & Bannister, 2005). The current study has identified differences between language pairs in the use of segmentation cues and provided an empirical foundation for future training studies. Participants in this study represent a diversity of language backgrounds. Since English is the lingua franca of the world, children and adults around the world are learning English both formally and informally. Although data were collected in the U.S., findings may potentially be generalized to ESL in other countries considering the representativeness of the sample. According to the U.S. Census 2011, there are over two million native Chinese speakers, over one million native Korean speakers, and over 34 million native Spanish speakers residing in the U.S. It is important to understand the acquisition process in L2 learners from these language backgrounds. Improved English proficiency may help new immigrants to become more competitive and productive in the work place.

Conclusion

This dissertation project examined the relative importance of various types of segmentation cues used in L2 speech comprehension. This project represents one of the first few attempts to improve the understanding of how L2 learners use a variety of cues in the speech signal to segment and comprehend continuous speech. Participants consisted of four language groups, including native English speakers and L2 learners of English with Korean, Mandarin or Spanish L1. In general, L2 learners appeared to rely more on sublexical cues than lexical cues in which segmental cues outweigh lexical knowledge and word-level semantics. However, L2 learners also showed native-like weighting of lexical cues in which both the English and L2 groups seemed to give more weight to sentence context than to lexical knowledge. Furthermore, while native Korean and Spanish L2 learners of English appeared to use both lexical and sublexical cues interactively, native English speakers showed a tendency to use cues independently. Finally, Spanish speakers' use of acoustic-phonetic cues may be influenced by phonological variation in L1. These results provided partial support for the Revised Framework for L2 Segmentation. In conclusion, this dissertation research demonstrated that the segmentation mechanism employed by L2 learners is similar, yet different, from those employed by native listeners. More importantly, although learners may have acquired complete control of the cues used in L2 segmentation, L1 knowledge remains active in nonnative listening,

Appendix A

Stimuli for Experiment 1

Real Word Context		Nonword Context		Prime	
SW	WS	SW	WS	SW	WS
Character	Consider	Manister	Dilicter	Register	Remember
Document	Department	Wattlement	Disbartent	Handicap	Heroic
Average	Advantage	Aivemige	Altamtige	Comedy	Cosmetic
Tournament	Opponent	Felement	Thezarent	Discipline	Dimension
Battery	Attorney	Bulnerty	Umony	Desperate	Distinguish
Element	Component	Aulement	Kisponent	Paradigm	Pacific
Elephant	Proponent	Adinent	Dristament	Vitamin	Vanilla
Minister	Disaster	Chukpenter	Mestander	Mediate	Mascara
Citizen	Horizon	Subizen	Fopaizen	Revenue	Recover
Cylinder	Surrender	Shanulter	Sekuter	Corridor	Consensus
Feasible	Ensemble	Wisiful	Thezamble	Tentative	Tomato
Massacre	Semester	Fownistre	Pinaster	Tangible	Tomorrow
Particle	Example	Vandiful	Vishemble	Penalty	Perspective
Messenger	Endeavour	Motander	Enparmer	Dedicate	Detergent
Bachelor	October	Yespiler	Alkaser	Passenger	Performance
Cinema	Umbrella	Tarila	Elboma	Calculate	Committee
Balcony	Spaghetti	Crinalty	Trydunny	Politic	Pathetic
Emperor	December	Thelimer	Gosember	Calibrate	Complexion
Generous	Tremendous	Tipersou	Profendous	Paradise	Petition
Algebra	Pajama	Elodra	Abela	Recipe	Republic

Appendix B

Stimuli for Experiment 2

Context						Target
/n/		/ŋ/		/s/		
Real word	Nonword	Real word	Nonword	Real word	nonword	
everyone	akluben	following	telewing	evidence	delbiens	already
discussion	pistaktion	anything	gelmeing	delicious	repamous	approval
protection	dispansion	beginning	puboting	religious	diligious	agreement
connection	simeksion	publishing	lekshering	universe	kiperous	attractive
direction	taneition	damaging	counmiding	enormous	konamous	agenda
determine	ritaimin	offering	avering	apprentice	ridrenous	exhibit
suggestion	fetestion	supporting	mirusing	dangerous	shanolous	exactly
origin	alorin	happening	disoothing	regardless	sofraness	example
expansion	egflocktion	promising	proruging	confidence	fagunous	equipment
bulletin	mulesin	editing	aliping	awareness	ilnantous	establish
election	elpertion	explaining	riksoulding	numerous	ductionous	initial
examine	uksaitin	responding	mikpasting	dynamics	tezenicks	incorrect
discipline	propertin	expecting	rispresing	benefits	subelents	interpret
veteran	baggeren	achieving	bowaging	ambulance	egmorance	ideally
genuine	bilemin	directing	bispekging	fabulous	zanihous	olympic
marathon	topersen	returning	manising	excellence	baritance	organic
Oxygen	salegion	relating	dishurting	appearance	abwelens	official
gasoline	marilin	applying	umeining	importance	igdalmons	opponent
transition	protresion	complaining	pomshinting	consensus	igstandous	unable
feminine	fybedin	copying	laizoing	tremendous	primaskus	unstable

Appendix C

Stimuli for Experiment 3

Context						Target
/n/		/ŋ/		/s/		
Related	Unrelated	Related	Unrelated	Related	Unrelated	
discipline	adoption	studying	betraying	syllabus	envelopes	mentor
religion	location	accepting	recording	faithfulness	unbalance	believe
protection	devotion	attacking	promising	resistance	audience	defense
donation	abdomen	budgeting	coloring	allowance	satellites	money
surgeon	creation	prescribing	retrieving	diagnose	delicious	doctor
dimension	everyone	advancing	borrowing	Celsius	miracles	degree
hurricane	suggestion	destroying	preferring	accidents	dialects	damage
construction	injection	creating	vomiting	architects	unconscious	design
illusion	horizon	enchanting	debating	performance	candidates	magic
infection	limousine	suffering	insulting	cancerous	ambitious	disease
aggression	proportion	murdering	responding	dangerous	happiness	monster
technician	cinnamon	programming	surrounding	instruments	benefits	machine
attraction	forgiven	sparkling	babbling	brilliance	principles	diamond
transaction	unicorn	bargaining	happening	economics	tolerance	market
champion	election	commanding	poisoning	apprentice	components	master
bulletin	destruction	reporting	becoming	announcements	consequence	message
subtraction	veteran	accounting	installing	estimates	ancestors	number
translation	deletion	pronouncing	sizzling	linguistics	investments	language
reception	collision	visiting	struggling	acceptance	semantics	welcome
guardian	petition	nurturing	surprising	generous	restaurants	mother

Appendix D

Stimuli for Experiment 4

Context	Sentence	After target
Related	Due to heavy snow storm and freezing rain, all airline companies CANcel	their flights.
Unrelated	An important client has arrived; the lawyers CANcel	all prior engagements.
Related	The spy was caught and offered a lot of money to reveal secrets but he would never beTRAY	his country.
Unrelated	Sometimes the researcher is required by law or university authorities to beTRAY	confidentiality.
Related	On Saturday morning, my children's favorite activity is to watch the carTOON	shows on TV.
Unrelated	After months of waiting, he is excited to see the new carTOON	illustrations.
Related	The visitor knocked but no one answered the door, so he left a MESSage	at the doorstep.
Unrelated	The mother checked her mail and found a MESSage	from her daughter.
Related	The jury reached a guilty verdict and the criminal was sentenced to jail for asSAULT	charges.
Unrelated	After searching online for several months, he decided to purchase asSAULT	rifles.
Related	Authors do not own the copyrights of their publication, instead, the copyrights beLONG	to the publisher.
Unrelated	Only faculty members who have formal support systems in place beLONG	to unions.
Related	Before Galileo made his discovery, people thought that earth was at the CENTer	of the galaxy.
Unrelated	At the construction site, workers are diligently shoveling materials into the CENTer	of the pile.
Related	Her parents are religious Christians, they taught her to beLIEVE	in God.
Unrelated	The game would be over soon, the coach told his players to beLIEVE	in themselves.
Related	Students all look forward to the arrival of their three-month long SUMmer	break.
Unrelated	If you want to apply for graduate school, now is the time to look for SUMmer	internships.
Related	Social anxiety sufferers often wind up unemployed or in jobs be LOW	their training level
Unrelated	After searching the house for several hours, he finally found his keys beLOW	the table.
Related	A fire broke out in the house, fortunately the only DAMage	was a burnt couch.
Unrelated	Scientists claimed that global warming may DAMage	health.
Related	They like all the rides in the amusement park, but they particularly enJOY	the rollercoaster rides.
Unrelated	Helped by the increasing cost of new treatments, the health-care industry will enJOY	steady growth.

Related	The cause of the extinction of dinosaurs is a controversial TOPic	of great debate.
Unrelated	The helpful journal editor has suggested a new TOPic	for the writer.
Related	The idea behind early voting is that, by making it easier for peo ple to vote, you will enSURE	that more people d o so.
Unrelated	The professor told his students to search the literature and enSURE	that no one has studied this topic before.
Related	After upgrading, there will be a new FEATure	in your phone.
Unrelated	She has a lot of admirers because her physical FEATures	are very attractive.
Related	With its enormous wings and breathing fire, the DRAGon	is intimidating.
Unrelated	There is no scientific evidence showing that any DRAGon	ever existed.
Related	Scientists have the moral obligation to publish their findings and inFORM	the public about their research.
Unrelated	In an arranged marriage, parents would choose a man for their daughter and inFORM	her of their decision as a surprise.
Related	Astronomers believe that Mars is the only possible PLANet	to have water.
Unrelated	The discovery channel is a great place to learn about the PLANet	earth.
Related	For the group project, students cannot choose their partners freely, instead, the teacher will asSIGN	them to their groups.
Unrelated	Although a dictator cannot command all of his people's obedience, he can asSIGN	them to their duty.
Related	When you turn on the new feature, you will see a red dot in the lower left hand CORner	of the screen.
Unrelated	The grand opening of the big supermarket took away customers from the CORner	shop.

Appendix E

Models estimating the effects of language group, stress location, and lexical knowledge on response latency and accuracy in the lexical decision task in Experiment 1.

Response Times. The intercept estimated English speakers' log RT in the baseline condition with nonword context and initial-stressed primes (Table 34). For brevity, only interactions among language groups, priming conditions, and one of the two segmentation cues were discussed. The interaction between language group and priming condition was not significant in all language groups, suggesting that the magnitude of priming effects of the L2 groups was similar to that of the native speakers. For the use of lexical knowledge (Figure 14), the English group showed a significant interaction between priming condition and context lexicality, demonstrating great priming effects (i.e. faster segmentation time) when context was a real word compared to when it was a nonword. In comparison to the native English group, all three L2 groups showed smaller priming effects when context was a real word (i.e. slower segmentation time evidenced by the positive estimates), although the interaction between priming condition and lexicality was only significant in the Korean and Mandarin groups. These results suggest that nonnative speakers rely less on lexical knowledge than native speakers.

For the use of stress cues (Figure 15), the English group showed a significant interaction between priming condition and stress location, demonstrating greater priming effects for WS words. In contrast, the priming condition by stress interaction was not significant in the Korean or Mandarin group. However, this interaction was marginally significant in the Spanish group, which showed smaller priming effect for WS words.

Accuracy. In the model predicting accuracy, there was no significant interaction involving priming conditions and language groups.

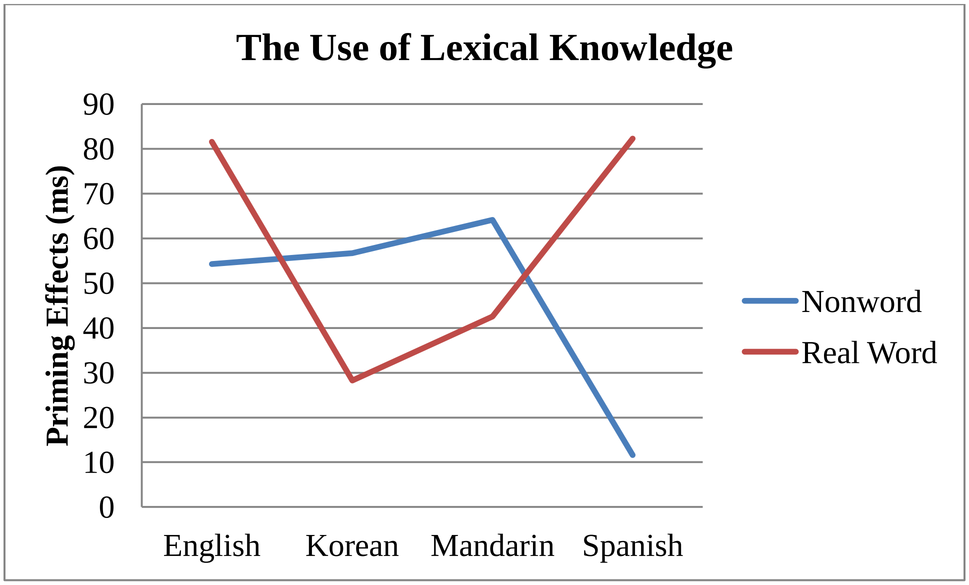

Figure 14 Language group differences in the use of lexical knowledge in segmentation

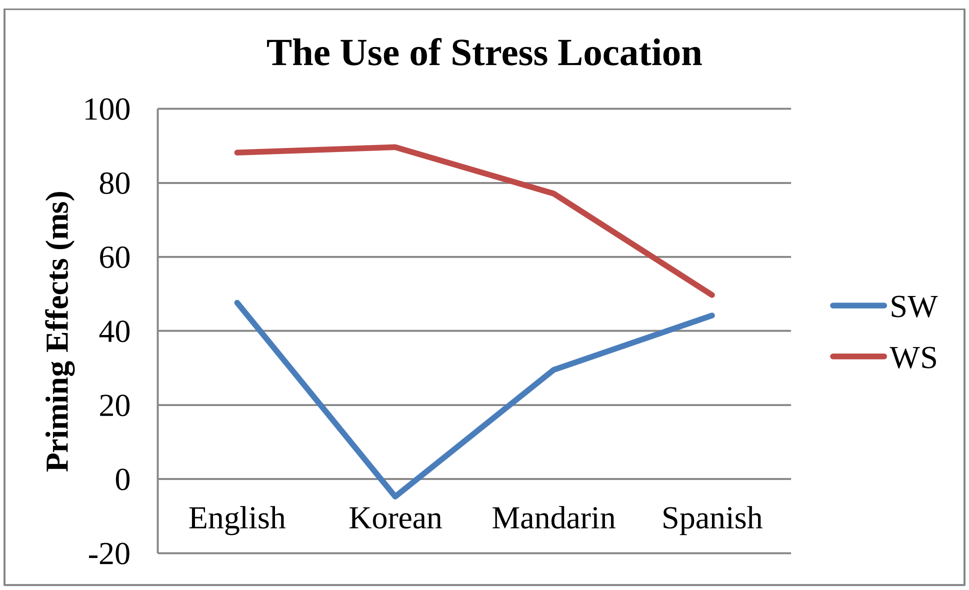

Figure 15 Language group differences in the use of stress location in segmentation.

Table 34 Mixed-effects linear model estimating log RT

Fixed Effects	Estimate	Std. Error	*t*-value	*p*MCMC
Intercept (English)	2.982	.065	47.75	.0001
English × Priming	-.012	.010	-1.21	.224
Korean × Priming	.008	.008	.96	.337
Mandarin × Priming	-.002	.008	-.27	.783
Spanish × Priming	.007	.008	.81	.406
English × Priming × Lexicality	-.016	.006	-2.51	.013
Korean × Priming × Lexicality	.015	.007	2.30	.023
Mandarin × Priming × Lexicality	.016	.007	2.34	.018
Spanish × Priming × Lexicality	.007	.007	1.06	.298
English × Priming × Stress	-.016	.006	-2.54	.009
Korean × Priming × Stress	.004	.007	.66	.512
Mandarin × Priming × Stress	.008	.007	1.16	.244
Spanish × Priming × Stress	.012	.007	1.73	.083

Appendix F

Models estimating the effects of language group, phonotactic cues, and lexical knowledge on response latency and accuracy in the word spotting task in Experiment 2.

Response Times. For the model predicting response latency (Table 35), the intercept estimated English speakers' log RT in the coda /n/ condition with nonword contexts. For brevity, only language group-related effects were reported. All three L2 groups showed significantly longer RT than the native group in spotting the target word when it was preceded by a nonword context. When the target word was preceded by a real word context, it also took the L2 groups significantly longer than the native group to identify the target word. Although all four groups of speakers showed faster RT when the target word was preceded by a real word context than when it was preceded by a nonword context, the difference in RT was larger in the native English speakers compared to the nonnative speakers (Figure 16).

Table 35 Mixed-effectss linear model estimating log RT

Fixed Effects	Estimate	Std. Error	*t*-value	*p*MCMC
Intercept (English)	3.101	.032	97.92	.0001
Korean	.039	.012	3.23	.0001
Mandarin	.030	.011	2.78	.0016
Spanish	.042	.011	3.73	.0002
Korean × coda /ŋ/	-.006	.005	-1.35	.177
Mandarin × coda /ŋ/	-.003	.005	-.68	.472
Spanish × coda /ŋ/	-.005	.005	-1.00	.313
Korean × coda /s/	-.002	.005	-.49	.623
Mandarin × coda /s/	-.001	.005	-.13	.879
Spanish × coda /s/	-.007	.005	-1.47	.144
Korean × Lexicality	.009	.004	2.57	.009

Mandarin × Lexicality	.009	.004	2.39	.018
Spanish × Lexicality	.008	.004	1.96	.0485

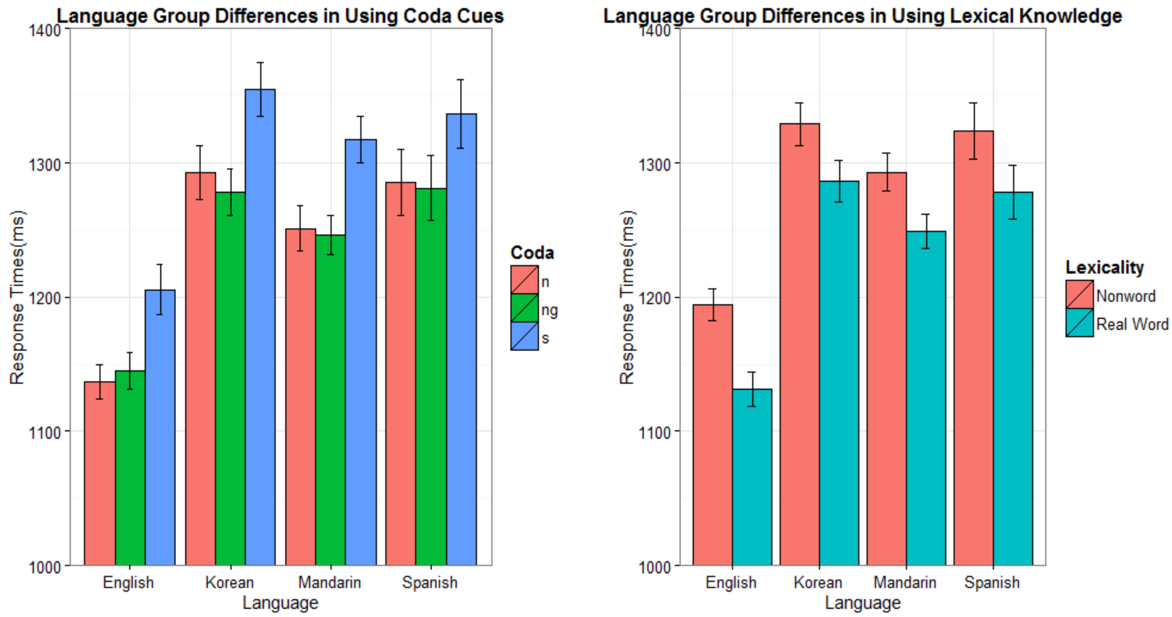

Figure 16 Language Group differences in the use of phonotactic cues and lexical knowledge (RT)

Accuracy. In the model predicting accuracy (Table 36), the intercept estimated English speakers' accuracy in the coda /n/ condition with nonword contexts. Mandarin speakers were marginally significant more accurate (evidenced by the positive estimates) than English speakers while Korean speakers were significantly more accurate than English speakers when the target word was preceded by a nonword context. As Figure 17 shows, English speakers were more accurate in the /ŋ/ condition than in the /n/ condition whereas the opposite was observed in Mandarin speakers. This language group by coda cue interaction was marginally significant. Both English and Mandarin speakers were more accurate in the /n/ condition than in the /s/ condition, although the difference in accuracy was significantly larger in the Mandarin group than in the English group. Finally, both Korean and English speakers were less accurate when the target word was preceded by a nonword context when it was preceded by a real word context; however, this difference in accuracy was significantly larger in the English group.

Table 36 Mixed-effects linear model estimating accuracy (log odds)

Fixed Effects	Estimate	Std. Error	*t*-value	*p*MCMC
Intercept (English)	2.541	.870	2.919	.0035
Korean	.579	.419	1.382	.167
Mandarin	.769	.421	1.828	.067
Spanish	.867	.451	1.921	.055
Korean × coda /ŋ/	.008	.436	.019	.985
Mandarin × coda /ŋ/	-.832	.456	-1.826	.067
Spanish × coda /ŋ/	-.675	.491	-1.374	.169
Korean × coda /s/	-.533	.361	-1.477	.139
Mandarin × coda /s/	-.879	.406	-2.164	.030
Spanish × coda /s/	-.052	.466	-.112	.911
Korean × Lexicality	-.714	.316	-2.26	.024
Mandarin × Lexicality	-.189	.334	-.567	.570
Spanish × Lexicality	-.600	.385	-1.557	.119

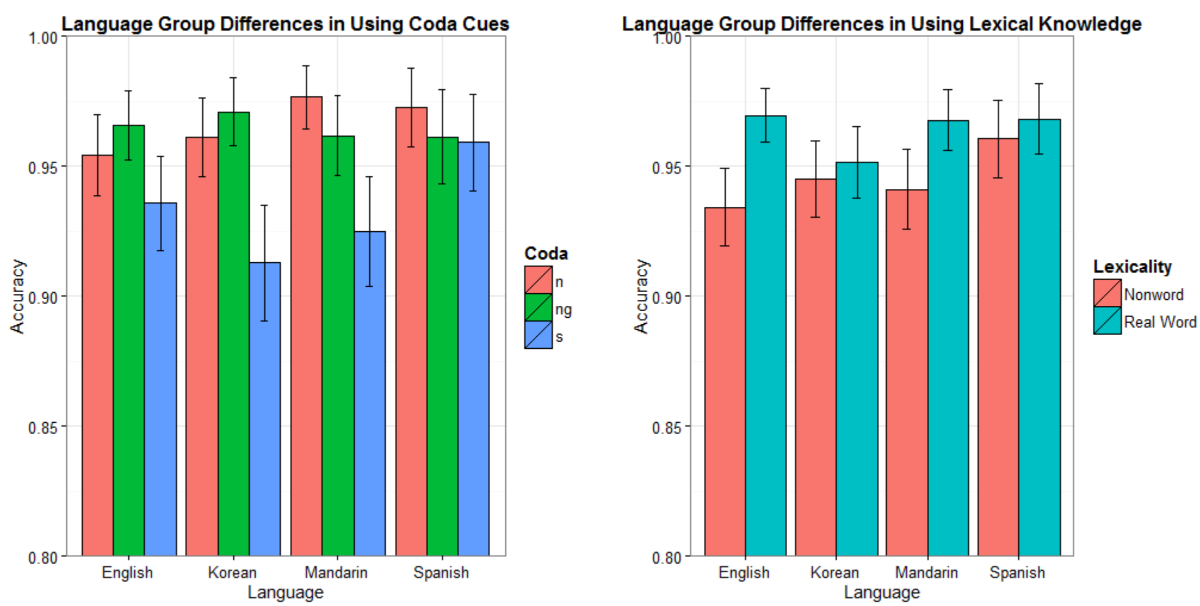

Figure 17 Language group differences in the use of coda cues and lexical knowledge (accuracy)

Appendix G

Models estimating the effects of language group, phonotactic cues, and semantic relatedness on response latency and accuracy in the word spotting task in Experiment 3.

Response Times. For the model predicting response latency (Table 37), the intercept estimated English speakers' log RT in the coda /n/ condition with semantically unrelated contexts. Only language group-related effects were reported. In the coda /n/ condition, there were significant effects of language group in which all three L2 groups identified the target words significantly slower than the native group. There was a significant interaction between English and Spanish for codas /n-s/. As Figure 18 shows, English speakers were faster than Spanish speakers in identifying target words preceded by /n/ as well as those preceded by /s/. However, this RT difference was larger for coda /s/ than for /n/. There was no significant interaction between language and semantic relatedness.

Table 37 Mixed-effectss linear model estimating log RT

Fixed Effects	Estimate	Std. Error	*t*-value	*p*MCMC
Intercept (English)	3.074	.029	104.39	<.0001
Korean	.041	.011	3.59	.0003
Mandarin	.034	.010	3.34	.0009
Spanish	.032	.010	3.01	.0026
Korean × coda /ŋ/	-.004	.005	-.76	.462
Mandarin × coda /ŋ/	.001	.005	.12	.895
Spanish × coda /ŋ/	.004	.005	.79	.428
Korean × coda /s/	.006	.005	1.33	.183
Mandarin × coda /s/	.008	.005	1.57	.116
Spanish × coda /s/	.012	.005	2.33	.018
Korean × Semantic relatedness	.002	.004	.44	.658

Mandarin × Semantic relatedness	-.004	.004	-.96	.344
Spanish × Semantic relatedness	-.004	.004	-.95	.347

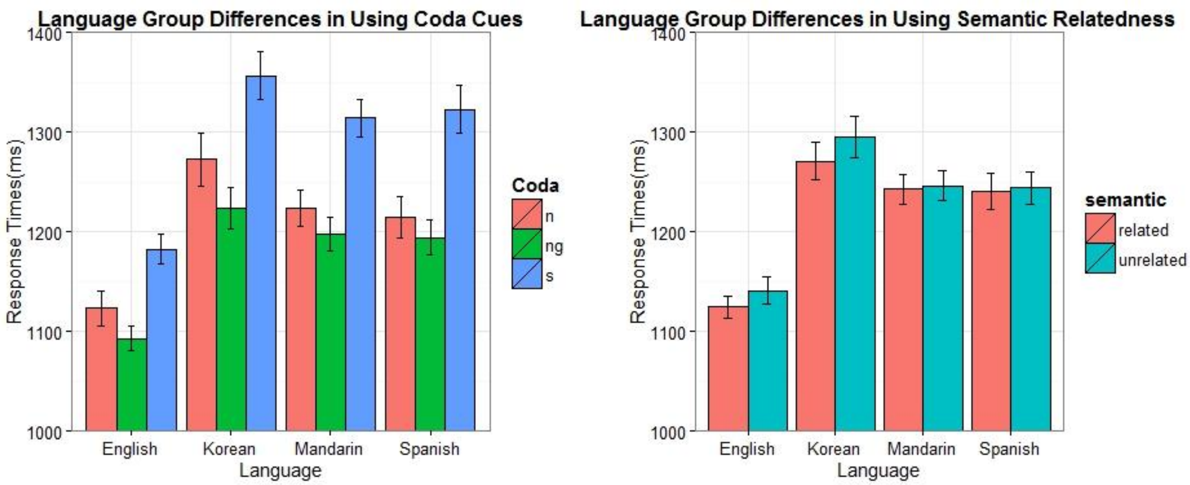

Figure 18 Language group differences in the use of phonotactic cues vs. semantic relatedness (RT)

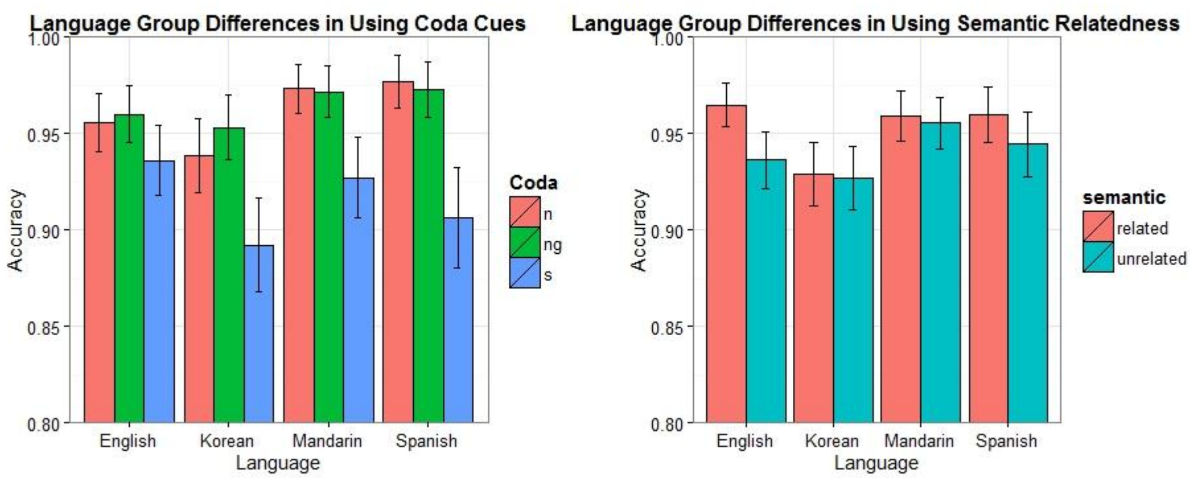

Figure 19 Language group differences in the use of phonotactic cues vs. semantic relatedness (accuracy)

Accuracy. For the model predicting response accuracy (Table 38), the intercept estimated English speakers' accuracy (in log odds) for the coda /n/ condition with semantically related contexts. There was a significant effect of the Korean group. For semantically related contexts with /n/ coda, Korean speakers were significantly less accurate than English speakers in word identification. There was a significant interaction between Spanish and English for codas /n-s/. As Figure 19 shows, English speakers were less accurate than Spanish speakers for coda /n/ whereas English speakers were more accurate than Spanish speakers for coda /s/. There was a significant interaction between Korean and English groups for semantic relatedness. English speakers were more accurate identifying target words preceded by semantically related context than those preceded by unrelated context; this semantic relatedness effect was not observed in the Korean group.

Table 38 Mixed-effectss linear model estimating accuracy (log odds)

Fixed Effects	Estimate	Std. Error	*t*-value	*p*MCMC
Intercept (English)	4.146	.778	5.327	<.0001
Korean	-.765	.385	-1.986	.047
Mandarin	.142	.420	.338	.735
Spanish	.532	.464	1.145	.252
Korean × coda /ŋ/	.190	.375	.506	.612
Mandarin × coda /ŋ/	-.171	.451	-.380	.704
Spanish × coda /ŋ/	-.292	.504	-.579	.562
Korean × coda /s/	-.239	.327	-.731	.465
Mandarin × coda /s/	-.674	.390	-1.726	.084
Spanish × coda /s/	-1.137	.429	-2.652	.008
Korean × Semantic relatedness	.602	.282	2.133	.033
Mandarin × Semantic relatedness	.533	.321	1.660	.097
Spanish × Semantic relatedness	.247	.337	.735	.462

Appendix H

Models estimating the effects of language group, sentence context, and stress location on response latency and accuracy in the lexical decision task in Experiment 4.

Response Times. For the model predicting response latency (Table 39), the intercept estimated English speakers' log RT in the baseline condition for unrelated sentence contexts and initial-stressed target words. There was a significant effect of priming for English speakers. There was a significant three-way interaction among language group (English vs. Spanish), priming condition, and sentence context. English speakers showed stronger priming effects for related sentence context (Figure 20) than for unrelated sentence context whereas the opposite pattern was observed in the Spanish group. There was a significant three-way interaction among language group (English vs. Spanish), priming condition, and stress location. Although both English and Spanish speakers showed stronger priming effects for initial-stressed target words than final-stressed target words, this difference in priming effects was significantly greater in the Spanish group (Figure 21).

Table 39 Mixed-effectss linear model estimating log RT

Fixed Effects	Estimate	Std. Error	*t*-value	*p*MCMC
Intercept (English)	2.959	.069	42.64	.0001
English × Priming	-.039	.016	-2.42	.016
Korean × Priming	.014	.025	.54	.551
Mandarin × Priming	-.019	.024	-.8	.442
Spanish × Priming	-.045	.026	-1.76	.092
English × Priming × Sentence context	-.014	.017	-.83	.431
Korean × Priming × Sentence context	-.0004	.018	-.02	.956
Mandarin × Priming × Sentence context	.029	.017	1.67	.102

Spanish × Priming × Sentence context	.040	.019	2.12	.042
Korean × Priming × Stress	.007	.017	.4	.68
Mandarin × Priming × Stress	.004	.018	.21	.840
Spanish × Priming × Stress	.020	.017	1.14	.253
Spanish × Priming × Stress	.070	.019	3.71	.0002

Accuracy. In the model predicting accuracy, there was no significant interaction involving priming conditions and language groups.

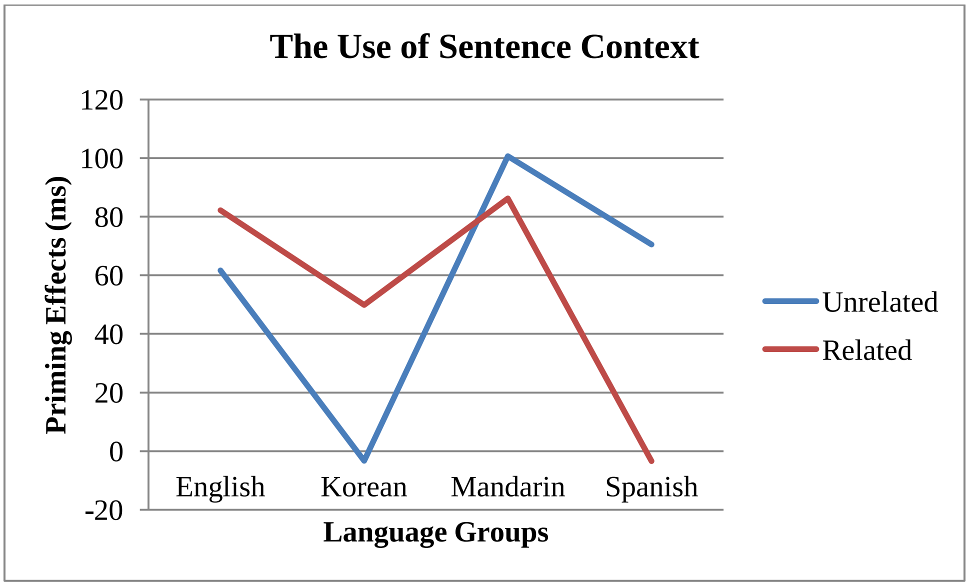

Figure 20 Language group differences in the use of sentence context in segmentation

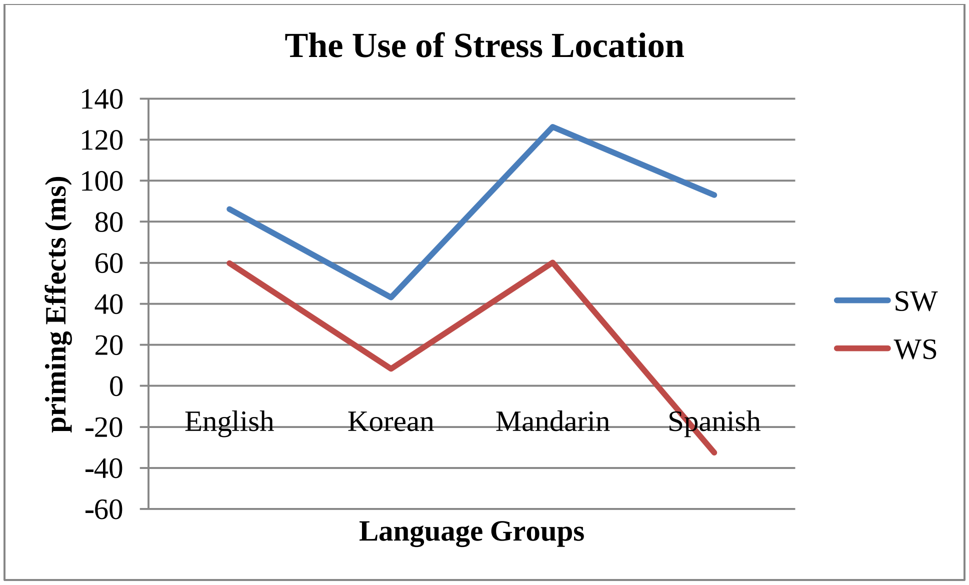

Figure 21 Language group differences in the use of stress location in segmentation.

Appendix I

Cloze Test

1. Skim the passage quickly to get the general meaning.
2. Read it carefully and supply only one word in EACH blank next to the item number. Contractions (e.g., don't) and possessives (e.g., John's) are one word. Try to supply as many missing words as you can.

Note: Spelling mistakes will not count against you as long as the scorer can read the word.

Man and His Progress

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

Up to 10,000 years ago, (12)_________ human beings lived by hunting and (13)_________. They also picked berries and fruit, (14)_______________ dug for various edible roots. Most (15)_________, the men were the hunters, and (16)___________ women acted as food gatherers. Since (17)_________ women were busy with the children, (18)________ men handled the tools. In a (19) _______ hand, a dead branch became a (20) __________ to knock down fruit or to (21)___________ for tasty roots. Sometimes, an animal (22)_______ served as a club, and a (23) _________ piece of stone, fitting comfortably into (24)_______ hand, could be used to break (25)________ or to throw at an animal. (26)_______ stone was chipped against another until (27) _______ had a sharp edge. The primitive (28)_______ who first thought of putting a (29) _____ stone at the end of a (30) _____made a brilliant discovery: he (31) _______ joined two things to make a (32)_______ useful tool, the spear. Flint, found (33)_____ many rocks, became a common cutting (34) _____ in the Paleolithic period of man's (35)______. Since no wood or bone tools (36)______ survived, we know of this man (37)_______ his stone implements, with which he (38)_____ kill animals, cut up the meat, (39)_______ scrape the skins, as well as (40)_______ pictures on the walls of the (41)__________ where he lived during the winter.

(42) _______ the warmer seasons, man wandered on (43)_______ steppes of Europe without a fixed (44)________, always foraging for food. Perhaps the (45) __________ carried

nuts and berries in shells (46)_______skins or even in light, woven (47) ________. Wherever they camped, the primitive people (48) ________ fires by striking flint for sparks (49)________ using dried seeds, moss, and rotten (50)_________ for tinder. With fires that he kindled himself, man could keep wild animals away and could cook those that he killed, as well as provide warmth and light for himself.

References

Al-jasser, F. (2008). The effect of teaching English phonotactics on the lexical segmentation of English as a foreign language. *System, 36*, 94-106.

Altenberg, E. P. (2005). The perception of word boundaries in a second language. *Second Language Research, 21*, 325-358.

Bachman, L F. (1982). The trait structure of cloze test scores. *TESOL Quarterly, 16*, 61-7.

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., Neely, J. H., Nelson, D. L., Simpson, G. B., & Treiman, R. (2007). The English lexicon project. *Behavior Research Method 39*, 445-459. Retrieved from http://elexicon.wustl.edu/.

Bardovi-Harlig, K. and Dörnyei, Z. (1998), Do language learners recognize pragmatic violations? Pragmatic versus grammatical awareness in instructed L2 learning. *TESOL Quarterly, 32*, 233–259.

Bates, D. M. (2005). Fitting linear mixed models in R: Using the lme4 package. *R News: The Newsletter of the R Project, 5*, 27-3.

Beckman, M., & Edwards, J. (1987). The phonological domains of final lengthening. *Journal of Acoustical Society of America, 81*, S67.

Boomershine, A. (2006). Perceiving and processing dialectal variation in Spanish: An exemplar theory approach. In T. L. Face & C. A. Klee (Eds.) *Selected proceedings of the 8th Hispanic Linguistics Symposium* (pp 58-72). Somerville, MA: Cascadilla Proceedings Project.

Borden, G. J., Harris, K. S., & Raphael, L. J. (2003). *Physiology, acoustics, and perception of speech*. Baltimore, MD: Lippincot Williams and Williams.

Bortfeld, H., Morgan, J. L., Golinkoff, R. M., & Rathbun, K. (2005). Mommy and me: Familiar names help launch babies into speech-stream segmentation. *Psychological Science, 16*, 298-304.

Chao, Y.-R. (1968). *A grammar of spoken Chinese.* Berkeley, CA: University of California Press.

Christie, W. M. (1974). Some cues for syllable juncture perception in English. *Journal of the Acoustical Society of America, 55*, 819-821.

Christophe, A., Peperkamp, S., Pallier, C. Block, E., & Mehler, J. (2004). Phonological phrase boundaries constrain lexical access I. Adult data. *Journal of Memory and Language, 51*, 523-547.

Church, K. (1983). Allophonic and phonotactic constraints are useful. International Joint Conference on Artificial Intelligence, Karlsruhe, West Germany.

Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics*, *27*, 3-42.

Clumeck, H. (1976). Patterns of soft palate movements in six languages. *Journal of Phonetics*, *4*, 337-351.

Cutler, A. (1997). The syllable’s role in the segmentation of stress languages. *Language and Cognitive Processes, 12*, 839-845.

Cutler, A. (2000). Listening to a second language through the ears of a first. *Interpreting, 5*, 1-23.

Cutler, A., & Carter, D. M. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech and Language*, *2*, 133-142.

Cutler, A. Mehler, J., Norris, D. G., & Segui, J. (1986). The syllable's differing role in the segmentation of French and English. *Journal of Memory and Language, 25,* 385-40.

Cutler, A., & Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance, 14,* 113-121.

Cutler, A., & Otake, T. (1994). Mora or phoneme? Further evidence for language-specific listening. *Journal of Memory and Language, 33,* 824-844.

Dalbor, J.B. **(**1997). *Spanish pronunciation: theory and practice 3rd edition*. New York, NY: Holt, Rinehart and Winston.

Davis, M. H., & Marslen-Wilson, W. D., & Gaskell, M. G. (2002). Leading up the lexical garden path: Segmentation and ambiguity in spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 28,* 218-244.

Davis, S. & Shin, S.-H. (1999). The syllable contact constraint in Korean: An optimality-theoretic analysis. *Journal of East Asian Linguistics, 8*, 285-312.

DeCasper, A. J. & Spence, M. J. (1986). Prenatal maternal speech influences newborns' perception of speech sounds. *Infant Behavior and Development, 9,* 133-150.

De Jong, K., Beckman, M. E., & Edwards, J. (1993). The interplay between prosodic structure and coarticulation. *Language and Speech, 36*, 197-212.

Dewey, G. (1950). *Relative frequency of English speech sounds, 2nd Edition*. Cambridge, MA: Harvard University Press.

Dilley, L., & McAuley, J. D. (2008). Distal prosodic context affects word segmentation and lexical processing. *Journal of Memory and Language*, *59*, 294-311.

Dilley, L., Mattys, S. L., & Vinke, L. (2010). Potent prosody: Comparing the effects of distal prosody, proximal prosody, and semantic context on word segmentation. *Journal of Memory and Language*, *63*, 274-294.

Duanmu, S. (2007). *The phonology of standard Chinese, second edition*. New York, NY: Oxford University Press.

Dupoux, E., Peperkamp, S., & Sebastian-Galles, N. (2001). A robust method to study stress "deafness". *Journal of Acoustical Society of America*, *110*, 1606-1618.

Eimas, P. D., Siqueland, E. R., Jusczyk, P. W., & Vigorrito, J. (1971). Speech perception in infants. *Science*, *171*, 303-306.

Endress, A. D., & Hauser, M. D. (2010). Word segmentation with universal prosodic cues. *Cognitive Psychology 61*, 177-199.

Eimas, P. D., Siqueland, E. R., Jusczyk, P. W., & Vigorrito, J. (1971). Speech perception in infants. *Science*, *171*, 303-306.

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

Felser, C., Roberts, L., Marinis, T., & Gross, R. (2003). The processing of ambiguous sentences by first and second language learners of English. *Applied Psycholinguistics, 24*, 453-489.

Field, J. (2008). Revising segmentation hypotheses in first and second language listening. *System, 36*, 35-51.

Flege, J. E. (1992). Speech learning in a second language. In C. Ferguson, L. Menn, & C. Stoel-Gammon (Eds.), *Phonological Development: Models, Research, and Implications* (pp. 565–604). Timonium, MD: York Press,

Flege, J.E. (1995). Second-language speech learning: Findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Theoretical and methodological issues* (pp. 233–273). Timonium, MD: York Press.

Flege, J. E. (2002). Interactions between the native and second language phonetic systems. In P. Burmeister, T. Piske, & Rohde, A. (Eds.), *An Integrated View of Language Development: Papers in Honor of Henning Wode*. Wissenschaftlicher Verlag Trier, Trier.

Fougeron, C., & Keating, P. (1997). Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustical Society of America*, *101*, 3728-374.

Fry, D. B. (1955). Duration and intensity as physical correlates of linguistic stress. *Journal of the Acoustical Society of America*, *27*, 765–768.

Fry, D. B. (1958). Experiments in the perception of stress. *Language and Speech, 1*, 126–152.

Furman, N., Goldberg, D., & Lusin, N. (2007). Enrollments in language other than

English in the United States institutions of higher education fall 2006. *The Modern Language Association*. Retrieved from http://www.mla.org/pdf/06enrollmentsurvey_final.pdf/

Gomez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science, 13,* 431–436.

Gow Jr., D. W., & Gordon, P. C. (1995). Lexical and prelexical influences on word segmentation: Evidence from priming. *Journal of Experimental Psychology: Human Perception and Performance, 21*, 344-359.

Guion, S. (2005). Knowledge of English word stress patterns in early and late Korean-English bilinguals. *Studies in Second Language Acquisition, 27*, 503–533.

Guion, S. G., Harada, T., & Clark, J .J. (2004). Early and late Spanish-English bilinguals' acquisition of English word stress patterns. *Bilingualism: Language and Cognition, 7*, 207-226.

Hayes, B. (1995). *Metrical stress theory: Principles and case studies*. Chicago: University of Chicago Press.

Hanulikova, A., McQueen, J. M., & Mitterer, H. (2010). Possible words and fixed stress in the segmentation of Slovak speech. *The Quarterly Journal of Experimental Psychology, 63*, 555-579.

Hanulikova, A., Mitterer, H., & McQueen, J. M. (2011). Effects of first and second language on segmentation of non-native speech. *Bilingualism: Language and Cognition, 14,* 506-521.

Hu, G., & Jiang, N. (2011). Semantic integration in listening comprehension in a second language: Evidence from cross-modal priming. In P. Trofimovich & K.

McDonough (Eds.) *Applying priming methods to L2 learning, teaching and research: Insights from psycholinguistics* (pp. 199-218) Amsterdam, The Netherlands: John Benjamins.

Iverson, P., Kuhl, P. K., Akahane-Yamada, R., Diesch, E., Tohkura, Y., Kettermann, A., & Siebert, C. (2003). A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition, 87*, B47-B57.

Iverson, P., Hazan, V., & Bannister, K. (2005). Phonetic training with acoustic cue manipulations: A comparison of methods for teaching English /r/-/l/ to Japanese adults. *Journal of Acoustical Society of America, 118*, 3267-3278

Jakobson, R. (1971). S*elected writings I*. Mouton: The Hague.

Jiang, N. (2002). Form-meaning mapping in vocabulary acquisition in a second language. *Studies in Second Language Acquisition, 24*, 617-637.

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

Jusczyk, P. W., & Aslin, R. N. (1995). Infants' detection of the sound patterns of words in fluent speech. *Cognitive Psychology, 29*, 1-23.

Jusczyk, P. W., Houston, D. M., & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology, 39*, 159-207.

Jun, S.-A. (2005). Korean intonational phonology and prosodic transcription. In: S.-A. Jun (Ed.) *Prosodic typology: The phonology of intonation and phrasing* (pp. 201-229). Oxford: University Press.

Karlsson, F. (1999). *Finnish: An essential grammar.* London, England: Routledge.

Kazanina, N., Phillips, C., & Idsardi, W. (2006). The influence of meaning on the perception of speech sounds. *Proceedings of the National Academy of Sciences, 103*, 11381-11386.

Kelly, J., & Local, J. (1989). *Doing phonology*. Manchester: Manchester University Press.

Kilborn, K. (1992). On-line integration of grammatical information in a second language. In R.Harris (Ed.), *Cognitive processing in bilinguals* (pp. 337–351). New York: Elsevier.

Kim, S., Broersma, M., & Cho, T. (2012). The use of prosodic cues in processing an unfamiliar language. *Studies in Second Language Acquisition, 34*, 415–444.

Klatt, D. H. (1973). Interaction between two factors that influence vowel duration. *Journal of Acoustical Society of America, 54*, 1102-1104.

Koda, K. (2004). *Insights into Second Language Reading*. NY: Cambridge University Press.

Krakow, R. A. (1989). The articulatory organization of syllables: A kinematic analysis of labial and velar gestures. *Doctoral Dissertation*, Yale University, New Haven, CT.

Krakow, R. A. (1999). Articulatory organization of syllables: A review". *Journal of Phonetics, 27*, 23-54.

Kroll , J. F. & Stewart , E. (1994). Category interference in translation and picture naming: evidence for asymmetric connections between bilingual memory representations . *Journal of Memory and Language , 33*, 149–174.

Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience 5*, 831-843.

Lehiste, I. (1960). An acoustic–phonetic study of internal open juncture. *Phonetica, 5*, 5–54.

Liberman, A.M., Cooper, F.S., Shankweiler, D.P., & Studdert-Kennedy, M. (1967). Perception of speech code. *Psychological Review, 74*, 431–461.

Lin, C. C. (2002). Nasal endings of Taiwan Mandarin: Production, perception, and linguistic change. Paper presented at the 35th International Conference on Sino-Tibetan Languages and Linguistics. Arizona State University, Tempe.

Lin, C. Y., Lukyanchenko, A., Winn, M. B., & Idsardi, W. (in preparation). Acoustic Cues to Perception of Word Stress by English, Mandarin and Russian Speakers.

Lin, C. Y., Wang, M., Idsardi, W. J., & Xu, Y. (under review). Stress perception in Mandarin-English and Korean-English bilinguals.

Local, J. (2003). Variable domains and variable relevance: interpreting phonetic exponents. *Journal of Phonetics, 31*, 321-339.

Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing, 19*, 1-36.

Lund, K., & Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, Instruments, & Computers 28*, 203-208.

Macpherson, I. R. (1975) *Spanish phonology: descriptive and historical*. Manchester: Manchester University Press.

Marian , V., Blumenfeld , H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech Language and Hearing Research*, *50***,** 940-967.

Marinis, T., Roberts, L., Felser, C., & Clahsen, H. (2005). Gaps in second language sentence processing. *Studies in Second Language Acquisition, 27*, 53-78.

Mattys, S. L. (2004). Stress versus coarticulation: Toward an integrated approach to explicit speech segmentation. *Journal of Experimental Psychology: Human Perception and Performance, 30*, 397-408.

Mattys, S. L., White, L., & Melhorn, J. F. (2005). Integration of multiple speech segmentation cues: A Hierarchical Framework. *Journal of Experimental Psychology: General, 134*, 477-50.

McQueen, J. M. (1998). Segmentation of continuous speech using phonotactics. *Journal of Memory and Language*, *39*, 21-46.

McQueen, J. M., Otake, T., & Cutler, A. (2001). Rhythmic cues and possible-word constraints in Japanese speech segmentation. *Journal of Memory and Language, 45*, 103-132.

Mok, P. (2008). On the syllable-timing of Cantonese and Beijing Mandarin. In *Proceedings of the 8th Phonetics Conference of China and the International Symposium on Phonetic Frontiers*.

Moon, C., Cooper, R. P., & Fifer, W. P. (1993). Two-day-olds prefer their native language. *Infant Behavior and Development, 16*, 495-50.

Nakatani, L. H., & Dukes, K. D. (1977). Locus of segmental cues for word juncture. *Journal of the Acoustical Society of America, 62*, 714-719.

Navarro, T. (1966). Estudios de fonologı´a espan˜ola [Studies of Spanish phonology]. New York: Las Ame´ricas.

Nazzi, T., Jusczyk, P., & Johnson, E. K. (2000). Language discrimination by Engilsh-learning 5-month-olds: Effects of Rhythm and Familiarity. *Journal of Memory and Language, 43*, 1-19.

Newman, R. S., Sawusch, J. R., & Wunnenberg T. (2011). Cues and cue interactions in segmenting words in fluent speech. *Journal of Memory and Language, 64*, 460-476.

Nguyen, T. A.-T. & Ingram, C. L. J. (2005). Vietnamese acquisition of English word stress. *TESOL Quarterly, 39*, 309–319.

Norris, D., McQueen, J. M., Cutler, A., & Butterfield, S. (1997). The possible-word constraint in the segmentation of continuous speech. *Cognitive Psychology, 34*, 191-243.

Norris, D., McQueen, J., Cutler, A., Butterfield, S., & Kearns, R. (2001). Language-universal constraints on speech segmentation. *Language and Cognitive Processes, 16*, 637-66.

Otake, T., Hatano, G., Cutler, A., & Mehler, J. (1993). Mora or syllable? Speech segmentation in Japanese. *Journal of Memory and Language, 32,* 258-278.

Otake, T., Hatano, G., & Yoneyama, K. (1996). Speech segmentation by Japanese speakers. In: Otake, T. and Cutler, A. (Eds.), *Phonological Structure and Language Processing: Cross-Linguistic Studies* (pp. 183-201). Mouton de Gruyter, Berlin,.

Peperkamp, S. & Dupoux, E. (2002). A typological study of stress "deafness". In C. Gussenhoven & N. Warner (Eds.) *Laboratory Phonology, 7* (pp. 203-240). Berlin: Mouton de Gruyter.

Peperkamp, S. (2004). Lexical exception in stress systems: Arguments from early language acquisition and adult speech perception. *Language, 80*, 98-126.

Perea, M., & Rosa, E. (2002). The effects of associated and semantic priming in the lexical decision task. *Psychological Research, 66*, 180-194.

Pierrehumbert, J., & Frisch. S. (1997). Synthesizing Allophonic Glottalization. In J.

Santen, R. Sproat, J. Olive, & J. Hirschberg, (Eds.), *Progress in Speech Synthesis* (pp. 9-26). New York: Springer,

Pineros, C. E. (2006). The phonology of nasal consonants in five Spanish dialects. In F. Martinez-Gil & S. Colina (Eds.) *Optimality-theoretic studies in Spanish phonology* (pp.146-171). Amsterdam, Netherlands: John Benjamins B.V.

Quené, H. (1992). Integration of acoustic-phonetic cues in word segmentation. In M. E. H. Schouten (Ed.), *The auditory processing of speech: From sounds to words* (pp. 349-356). Berlin: Mouton de Gruyter.

R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org. 2008.

Roberts, L., & Felser, C. (2011). Plausibility and recovery from garden paths in second language sentence processing. *Applied Psycholinguistics, 32,* 299-331.

Sanders, L. D., Neville, H. J., & Woldroff, M. G. (2002). Speech segmentation by native and non-native speakers: The use of lexical, syntactic, and stress-pattern cues. *Journal of Speech, Language, and Hearing Research, 45*, 519-53.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274*, 1926-1928.

Sebastian-Galles, N., Dupoux, E., Segui, J., & Mehler, J. (1992). Contrasting syllabic effects in Catalan and Spanish. *Journal of Memory and Language, 31*, 18-32.

Shatzman, K. B., & McQueen, J. M. (2006). Segment duration as a cue to word boundaries in spoken-word recognition. *Perception & Psychophysics, 68*, 1-16.

Shen, X. S. (1993). Relative duration as a perceptual cue to stress in Mandarin. *Language and Speech, 36*, 415-433.

Siptar, P., & Torkenczy, M. (2000). *The phonology of Hungarian*. Oxford, UK: Oxford University Press.

Sohn, H-M. (1999). *The Korean language*. Cambridge, UK: Cambridge University Press.

Sole, M.-J. (1992). Phonetic and phonological processes: The case of nasalization. *Language and Speech, 35*, 29-43.

Stockwell, R. P., & Bowen, J. D. (1965). *The sounds of English and Spanish*. Chicago, IL: The University of Chicago Press.

Stockewell, R. P., Bowen, J. D., & Silva-Fuenzalida (1956). Spanish juncture and intonation. *Language, 32*, 641-665.

Storkel, H. L. (2001). Learning new words: Phonotactic probability in language development. *Journal of Speech, Language, and Hearing Research, 44*, 1321-1337.

Tanowitz, J., & Beddor, P. S. (1997). Temporal characteristics of coarticulatory vowel nasalization in English. *Journal of Acoustical Society of America, 101*, 3194(A).

Trail, A (1985). *Phonetic and phonological studies of !Xóõ Bushman*. Hamburg: Helmut Buske.

Tremblay, A., Coughlin, C. E., Bahler, C., & Gaillard, S. (2012). Differential contribution of prosodic cues in the native and non-native segmentation of French speech. *Journal of Laboratory Phonology, 3,* 385–423.

Thiessen, E. D., & Saffran, J. R. (2003). When cues collide: Use of stress and statistical cues to word boundaries by 7- to 9-month-old infants. *Developmental Psychology*, *39*, 706-716.

Tyler, M. D., & Cutler, A. (2009). Cross-language differences in cue use for speech segmentation. *Journal of the Acoustical Society of America, 126*, 367–376.

Umeda, N. (1975). Vowel duration in American English. *Journal of the Acoustical Society of America, 58*, 434–445.

van Heuven, V. J., & Hagman, P. J. (1988). Lexical statistics and spoken word recognition in Dutch. In P. Coopmans, & A. Hulk (Eds.), *Linguistics in the Netherlands, 1988* (pp.59–68). Dordrecht: Foris.

Vivevitch, M. S. & Luce, P. A. (2004). A web-based interface to calculate phonotactic probability for words and nonwords in English. *Behavior Research Methods, Instruments, & Computers, 36*, 481-487.

Vroomen, J., Tuomainen, J., & de Gelder, B (1998). The roles of word stress and vowel harmony in speech segmentation. *Journal of Memory and Language, 133-149.*

Vroomen, J., van Zon, M., & de Gelder, B. (1996). Cues to speech segmentation: Evidence from juncture misperceptions and word spotting. *Memory & Cognition, 24*, 744-755.

Wayland, R., Guion, S. G., Landfair, D., & Li, B. (2006). Native Thai speakers' acquisition of English word stress. *Journal of Psycholinguistics Research, 35*, 285-304.

Weber, A. (2000). The role of phonotatics in the segmentation of native and nonnative continuous speech. In *Proceedings of Workshop on Spoken Word Access Processes* (pp. 143-146). Nijmegen, The Netherlands.

Weber, A., & Cutler, A. (2006). First-language phonotactics in second-language listening. *Journal of the Acoustical Society of America, 119*, 597-607.

White, L., Melhorn, J. F., & Mattys, S. L. (2010). Segmentation by lexical subtraction in Hungarian speakers of second-language English. *The Quarterly Journal of Experimental Psychology, 63*, 544-554.

Whitley, M. S. (2002), *Spanish/English contrasts: a course in Spanish linguistics* (2nd ed.), Washington D.C.: Georgetown University Press.

Xu, Y.,& Liu, F. (2006). Tonal alignment, syllable structure and coarticulation: toward an integrated model. *Italian Journal of Linguistics,18*, 125–159.

Yip, M. C., (2004). Possible-word constraints in Cantonese speech segmentation. *Journal of Psycholingusitic Research*, *33*, 165-173.