The Role of Alternation in Phonological Relationships

A Dissertation Presented

by

Yu-an Lu

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

Doctor of Philosophy

in

Linguistics

Stony Brook University

August 2012

Stony Brook University

The Graduate School

Yu-an Lu

We, the dissertation committee for the above candidate for the

Doctor of Philosophy degree, hereby recommend

acceptance of this dissertation.

Ellen Broselow – Dissertation Advisor
Professor, Department of Linguistics

José Elías-Ulloa – Chairperson of Defense
Assistant Professor, Hispanic Languages and Literature

Marie K. Huffman
Associate Professor, Department of Linguistics

Arthur Samuel
Professor, Department of Psychology

Kathleen Currie Hall
Assistant Professor, City University of New York

This dissertation is accepted by the Graduate School

Charles Taber
Interim Dean of the Graduate School

Abstract of the Dissertation

The Role of Alternation in Phonological Relationships

by

Yu-an Lu

Doctor of Philosophy

in

Linguistics

Stony Brook University

2012

The concept of phonological relationships has been central in most, if not all, theories of phonology. The goal of this dissertation is to determine the contributions of two factors, *distribution* and *alternation*, in leading speakers to group sounds as members of the same category. Using previously established methods of testing speakers' perception and processing of sounds—similarity ratings, discrimination on a continuum, and semantic priming—I investigate the processing of coronal fricatives in three different languages: (i) English, in which the contrast between *s* and *sh* may signal differences in meaning (as in *see* vs. *she*), though the two sounds participate in limited morphological alternations as in *press/pressure*; (ii) Korean, in which *s* and *sh* are in complementary distribution and participate in regular and productive morphological alternations; and (iii) Mandarin, in which *s* and *sh* are in complementary distribution but do not participate in allomorphic alternations due to Mandarin's lack of affixation and its phonotactic

restrictions. The relationship between *s* and *sh* in Mandarin, due to the conflicting evidence from distribution and alternation, has been a matter of controversy. The results from the similarity rating experiment showed that both the Mandarin and English speakers rated *s* vs. *sh* as more different than did Korean speakers, suggesting that the Mandarin speakers, who have access only to *distributional* evidence, are less likely to treat *s/sh* as members of a single category than the Korean speakers, who are exposed to evidence from both *distribution* and *morphological alternation*. Furthermore, the judgments from the speakers of all three languages varied in different vowel contexts, suggesting that the assignment of two sounds as members of the same or separate categories is not necessarily absolute. These findings suggest that multiple factors contribute to the formation of phoneme categories and that phonological relationships are *gradient* rather than categorical.

To my grandma in heaven

獻給來不及知道我會唸書的阿媽
盧陳昏

Table of Contents

List of Figures

List of Tables

Acknowledgments

There were countless blessings in the past six years given to me through the people around me that helped me make completing this dissertation even possible. I am just so grateful to be at this stage where I can count some of my blessings. My deepest gratitude goes to my advisor, Ellen Broselow, for her guidance through every step of the dissertation writing. The many hours she spent discussing with me and reading through my drafts impacted me not just in the direction of the dissertation but also in shaping me into a better researcher. Her patience in guiding me is what I appreciated the most. I will always be indebted to her for training me into an independent researcher. I would also like to thank my committee members, Marie Huffman, Arthur Samuel, José Elías-Ulloa, and Kathleen Curie Hall. Marie's insightful comments had been helpful to me ever since the idea of the dissertation was generated from her seminar. Arty's training on statistics gave me the most valuable tool for my research. Beto's careful examination on my stimuli always brought me back to look at the basics. And I can't thank Kathleen enough for her thorough comments and suggestions on my dissertation. It's truly a blessing to have someone who knows the topic so well and shares her thoughts so generously. I am also thankful for Christina Bethin for her earlier comments and input on the dissertation proposal. I have also benefitted from the feedback and discussions with the members in the Stony Brook linguistic community: Mark Aronoff, Bob Hoberman, Lori Repetti, Jiwon Hwang, Sophia Kao, and Jennifer Park. Special thanks go to John Drury who helped me in exploring future projects for my dissertation.

A large portion of the dissertation involved experiment running. There are several individuals that I am so grateful for for helping me complete these experiments. I would like to thank Yuwen Lai for welcoming me to her lab and making running experiments possible in Taiwan. I also thank Yunju Suh for being my Korean brain on the dissertation. Without her, I wouldn't even know how to start designing the Korean semantic priming experiment when I only know how to say 'how are you' and 'where is the bathroom' in Korean. Hijo Kang had been my agent for recruiting Korean participants. Hijo brought in 60 amazing Korean participants who gave perfect data for my experiments. Mark Lindsay's wonderful voice provided me just the right stimuli for the English experiments. Special thanks go to all of my research participants for providing me the invaluable data.

I would also like to thank all of the linguistics faculty and graduate students for the wonderful six years. To Dan Finer, thank you for the guidance on my QP and in several seminars of yours. To John Bailyn, you rock my syntax world! To Richard Larson, thank you for giving me the confidence I needed at the right moment. To Marianne Catalano, thank you for always thinking of me whenever a teaching opportunity came up. To Andrei Antonenko, it has been fun working with you on various events—we are the best organizers ever! To Ling Wang and Sergei Srednyak, I'll always remember the laughs we shared. To Hisako Takahasi, thank you for being there for my ups and downs. My thanks go especially to Miran Kim, whom I know will support me where ever I am. To Chih-hsiang Shu, Young-ran An, Hyun-ju Kim, Susana Huidobro, Carlos de Cuba, Yukiko Asano, Sara Catlin, Ivana Mitrović, and Poppy Slocum, you enriched my life at different points in the past six years. To Sandra Brennan and Michelle Carbone, thank you for all the administrative help you provided.

I am also blessed with so many people in my graduate career who provided me different opportunities along the way and gave just the right dosage of encouragement when I needed it. To Hui-chuan J. Huang, thank you for cheering for me, having confidence in me, and giving me solid training before I came to Stony Brook. To Henry Yungli Chang, thank you for supporting me financially for my fieldwork in Taiwan that directly contributed to my two QPs. To Hui-chuan Hsu, Chiu-yu Tseng, Chen-Sheng Liu, and Yi-Ting Chen, thank you for encouraging me when I search for the next stop in my career. To Kie Zuraw and Edith Aldridge, thank you for giving me the opportunity to learn from you. Finally, to Chia-Yin Pan, Ailing Liao, Shawn Chang, and Kate Shen, thank you for always keeping me in your prayers.

我要謝謝所有的家人，感謝你們不時的關心問候，感謝你們的禱告，讓我在國外唸書這段時間不至匱乏。謝謝我的阿媽劉黃巧雲，您的禱告大大有力，上帝給我的恩典遠遠超過我能想像。感謝我的爸爸盧振哲、媽媽劉秋旻，無條件支持我，給我足夠的自由追求想做的事，知道我想家，幾乎每年都長途跋涉來看我。謝謝我的姐姐盧郁文、盧郁心、姐夫 Sung Lee、外甥李迦勒，讓我在紐約有最親的家人可以依靠。To Uncle James, Aunty Sue, Bihling, and Justin Wu, thank you for welcoming me to your family，你們給了我在紐約的第二個家。謝謝我的公婆，邱垂榮、吳玉女，以及哥哥邱振訓，您們是我最棒的啦啦隊!還有我的先生邱振豪，謝謝你的支持、陪伴，你讓這一切的努力都值得。最後，感謝天父的看顧，我回想這一路，恩典滿溢。

This work was generously supported by NSF grant BCS-07460227 to Ellen Broselow, Marie Huffman & Nancy Squires, and Chiang-Ching-Kuo Foundation Dissertation Fellowship (蔣經國國際學術交流基金會博士論文獎學金) to the author. The many errors that remain in this work are among the few things for which I can claim complete credit.

Chapter 1 Introduction

Most phonologists have recognized the need to distinguish sound differences (or feature differences) that are contrastive from those that are not. The traditional definition of contrast has relied heavily on the *distribution* of two given sounds. If two sounds occur in the same environment and substituting one for the other may signal lexical differences, then the sounds are considered to be in *contrast*, belonging to discrete *phoneme* categories (e.g., Swadesh 1934; Bloch 1948, 1950; Harris 1951; Moulton 1962; Trubetzkoy 1969; Dixon 1970; Vennemann 1971; Banksira 2000; Hualde 2004; Bullock & Gerfen 2005; Hall 2009). The classic test for contrast is by way of the *minimal pair* test: a minimal pair (e.g., [**s**i] 'see' vs. [**ʃ**i] 'she') consists of two forms with distinct meanings that differ by only one segment found in *the same phonological context* (Trubetzkoy 1969). On the other hand, two *phonetically similar* sounds in *complementary distribution*, where the choice of one vs. the other is predictable from the environment, are considered to be non-contrastive allophones, members of the same phoneme category (Trubetzkoy 1969: 46). An often-used example to demonstrate variants of the same phoneme in complementary distribution is through a superhero analogy. Superman and his alter ego Clark Kent differ in superficial aspects, but they are identical in terms of height, hair color, etc., demonstrating a property analogous to phonetic similarity. Furthermore, because they are variants of a single person, they can never be found in the same environment, demonstrating the property of complementary distribution. However, as Hall (2009) has demonstrated, predictability from context is not necessarily all-or-nothing. Hall argues that the difference between contrastive phonemes and allophonic variants is therefore gradient rather than categorical based on distributional predictability.

Distributional restrictions may serve as static constraints on lexical items. However, when the choice of variant is dependent on context, we frequently also find *morphological alternations*; that is, morpho-syntactic processes (e.g., affixation, compounding) that change the phonological context in which a sound appears may cause a single morpheme to exhibit different variants in different contexts. Therefore, alternation may also be a criterion for determining the relationships among sounds (Baudouin de Courtenay 1972; Anderson 1985). For example, the last sound in the morpheme *ride* [ɹaɪ**d**] alternates with [ɾ] when the same morpheme is suffixed with *-er,* which provides a tapping environment (i.e., intervocalic position preceding an unstressed syllable), as in *rider* [ɹáɪ**ɾ**-ɚ]. Given the fact that this kind of morphological alternation provides evidence for considering [d] and [ɾ] as variants of the same phoneme category, alternation is not always an indicator of phoneme membership, since the contrast between separate phonemes may be neutralized in some contexts created by morphological processes (Trubetzkoy 1969). Dutch final devoicing is such a case where the contrast between voiced and voiceless obstruents is neutralized in word final position, causing a single morpheme to be realized with either a voiced or voiceless obstruent depending on context (e.g., /not/ 'nut', sg. [no**t**] vs. pl. [no**t**-ə]; /nod/ 'necessity', sg. [no**d**] vs. pl. [no**d**-ə]).[1]

Other criteria used to determine whether sounds are contrastive or not include native speaker intuition (Swadesh 1934)—whether native speakers of a language recognize one sound as different from another—and orthography (Chao 1934)—whether the difference between two sounds is represented in the orthographic system of a language. However, in many cases only a subset of these factors is applicable, and in some cases different criteria may yield conflicting results. Then what factors cause language learners to assign sounds to different categories or to a single category?

The goal of this dissertation is to determine the contributions of *predictability* and *alternation* in leading speakers to group sounds as members of the same phoneme category. I investigate the processing of coronal fricatives, *s* and *sh*, in three different languages (English, Korean, and Mandarin) in which the two sounds participate in different types of relationships. The relationship between *s* and *sh* in Mandarin has been a matter of long-standing controversy. The choice of Mandarin *s* vs. *sh* is largely predictable from the environment: the palatal *sh* occurs before high-front vowels [i, y] or glides [j, ɥ], and *s* occurs in the context of non-high-front vowels or glides. However, despite this predictable distribution, the two sounds *do not*

[1] Note that the case of neutralization in Dutch has been shown to be incomplete (see Warner et al. 2004 and the references in there).

participate in morphological alternations, due to Mandarin's lack of affixation and its stringent restrictions on possible syllable structures (e.g., Chao 1931, 1934; Hartman 1944; Cheng 1968, 1973; Lin 1989; Chiang 1992; Wu 1994; Yip 1996; Duanmu 2007). In other words, there are no morphological conditions in which we see *s* and *sh* alternating to support the relatedness of the two sounds in Mandarin (section 1.3.3). The distributional predictability of these sounds has led some researchers to argue that *s* and *sh* should be considered variants of the same category, while the lack of alternations has led others to argue that *s* and *sh* should be considered separate categories. To help settle this question, I compare Mandarin speakers' treatment of these sounds with the treatment of parallel sounds by speakers of two other languages. English differs from Mandarin in that the distribution of *s* and *sh* is not predictable from the environment, and the difference between the sounds may be used to signal lexical differences (although the sounds do participate in limited alternations, as in *press*/*pressure*, section 1.3.1). Korean, on the other hand, shares with Mandarin the predictable distribution of *s* and *sh*, but differs from Mandarin in that these sounds also participate in regular and productive morphological alternations (section 1.3.2). Note that although I will use the spelling *sh* to represent the post-alveolar fricatives in English ([ʃ]) and in Korean and Mandarin ([ɕ]), these sounds are phonetically different. Ladefoged and Maddieson (1996) describe the major difference between [ʃ] and [ɕ] as "in the degree of raising of the front of the tongue" (1996: 153), adding that "ʃ has added lip rounding or protrusion" (1996: 148).

I will report on three experiments designed to investigate the behavior of speakers of these three languages with respect to the *s/sh* sound difference. If distributional predictability alone is sufficient to cause learners to assign two sounds to a single category, we would expect Mandarin and Korean speakers to treat these sounds similarly, in contrast to English speakers, who should assign them to different categories. On the other hand, if alternations are a necessary criterion for causing learners to analyze two sounds as members of a single phoneme category, we would expect Mandarin speakers to pattern with English speakers, in contrast to Korean speakers, for whom *s/sh* regularly alternate.

Three methods were used to probe the way in which speakers analyze the relationship between the two sounds: similarity ratings (Chapter 2), discrimination on a continuum (Chapter 3), and semantic priming (Chapter 4). These probes were chosen based on previous studies, discussed in section 1.2, showing that variants of the same phoneme are processed differently than contrastive phonemes. The results, taken together, suggest that there is not a simple answer to the question of whether two sounds are members of the same category, that *multiple factors* contribute in deciding category membership, and that phonological relationships are *gradient* rather than categorical (Goldsmith 1995; Hall 2009). In the similarity rating experiment, both the Mandarin and English speakers rated *s* vs. *sh* as more different than did Korean speakers. These results suggest that the Mandarin speakers, who have access only to *distributional* evidence for a relationship between *s* and *sh*, are less likely to treat these sounds as members of a single category than the Korean speakers, who are exposed to evidence from both *distribution* and *morphological alternation* for a *s/sh* relationship. However, in the discrimination experiment, in which speakers of all three language groups were presented with pairs of sounds varying by equal intervals on a *s-sh* continuum, language background did not correlate with significant differences in the ability to discriminate the two sounds (see discussion in Chapter 3). Finally, in a semantic priming experiment, the English speakers' results showed no evidence for a priming relationship between *s* and *sh* while for the Mandarin and Korean speakers, the results supported a priming relationship between *s* and *sh*. These results suggest that the Mandarin and Korean speakers, who are both exposed to distributional evidence supporting a *s/sh* relationship, showed similar effects of the *s/sh* difference in lexical processing (see discussion in Chapter 4).

The fact that the results of these experiments did not yield a uniform pattern—the Mandarin group patterned with the English group in their similarity judgments, but patterned with the Korean group in the semantic priming experiment—suggests that the assignment of these sounds to phoneme categories is gradient rather than absolute, as argued by Hall (2009). Furthermore, in the similarity rating experiment, judgments from the speakers of *all three languages* varied in different vowel contexts, suggesting that the assignment of these sounds to phoneme categories varies according to the *environment* in which *s* and *sh* reside (see discussion in Chapter 2). In Chapter 5 I discuss the implications of these results for the assumptions made by different phonological theories. The following sections briefly lay out the extent to which the role of alternation, as opposed to predictable distribution, has been used in different phonological analyses to group sounds as members of the same category (section 1.1), as well as some psycholinguistic studies suggesting that morphological alternations may affect speakers' perception of sounds (section 1.2). Section 1.3 provides a summary of the status of the coronal fricatives in the three languages investigated in the dissertation. The comparison of the results from the experimental probes in the three languages enables us to access the relative contributions of predictability and alternation.

1.1 The role of alternation vs. distribution in defining phonological relationships

Although morphological alternations often reinforce the role of phonological context in determining the occurrence of particular sounds, morphological alternation and other criteria are “usually used only in conjunction with the primary criteria [distribution and lexical distinction] in cases of conflict or uncertainty” (Hall 2009: 2). Morphological alternation has sometimes even been argued to be irrelevant to phonological analyses (e.g., Trager 1934; Hockett 1942). For example, Trager (1934: 340) argued that alternation “does not properly concern us in a purely phonemic study.” Silverman (2006), on the other hand, argues that “learning allophonic relations is dependent upon learning allomorphic relations” (Silverman 2006: 26) and that “the only way sounds can be allophonically related is if they alternate with each other” (Silverman 2006: 88). One of Silverman’s arguments for *alternation* as the *only* diagnostic, and his rejection of distributional predictability, comes from Akan, a language of Ghana. In Akan, the choice of dorsal [k] vs. palatal [tç] is *largely predictable* from the phonological environment. The palatal [tç] occurs before non-low front vowels ([i, ɪ, e, ɛ]), as in (a-c) in Table 1-1, while the dorsal [k] occurs before other vowels ([u, ʊ, o, ɔ, ɑ]), as in (d-f).

Table 1-1 Complementary distribution of [k] and [tɕ] in Akan

a.	tɕɛ	‘divide’
b.	tɕim	‘umbrella’
c.	ɔtɕe	‘river’
d.	kɔʔ	‘go’
e.	kun	‘kill’
f.	ka	‘to bite’

In face of the predictable distribution, some phonologists might posit a single underlying representation for the two sounds, and derive surface [tɕ] via a palatalization rule. However, Silverman argues that a reduplication process in Akan provides evidence against deriving these two sounds from a single category. As shown in Table 1-2, the reduplicative prefix consists of a copy of the initial consonant of the base, followed by a vowel that is high but shares the backness, roundness, and tenseness of the base vowel (e.g., [e]↔[i]; [ɛ]↔[ɪ]; [o]↔[u]). Because there is no high back unrounded vowel, the raised correspondent of the nonround back vowel [a] is [ɪ].

Table 1-2 Reduplication in Akan

a.	si+siʔ		‘stand’
b.	su+soʔ		‘seize’
c.	kɪ+kaʔ	*tɕɪ+kaʔ	‘bite’

In (c), where the base vowel is [a], the dorsal consonant is placed before a front vowel [ɪ] in the reduplicative prefix, the context where the palatalization rule *should* apply. Silverman thus argues that Akan speakers, who are exposed only to evidence from static distribution, do not make the generalization that [k] and [tɕ] are derived from the same underlying representation (Silverman 2006: 104).[2]

However, McCarthy & Prince (1995) provide an analysis of these facts within the framework of Optimality Theory (Prince & Smolensky 1993) which allows general distributional constraints to be violated only in reduplicative affixes by appealing to the special nature of the relationship between *base* and *reduplicant*. The case of Akan is analyzed using constraints that demand identity between the place features of the consonant in the *base* (e.g., the dorsal consonant in [**k**aʔ]) and the place features of the copied consonant in the *reduplicant* ([**k**ɪ+kaʔ]). These base-reduplicant correspondence constraints outrank the general structural well-formedness requirements which govern the distribution of [k] and [tɕ]. This analysis does not require a commitment to deriving [k] and [tɕ] from the same underlying segment, but at the same time takes into account the distributional relationship between the two sounds.

As the above discussion illustrates, agreement on a set of criteria determining phonological relationships has not been reached. The next section reviews related psycholinguistic studies on the perception and processing of sounds with different phonological relationships.

[2] See also a review of Silverman's arguments in Dunbar & Idsardi (2010).

1.2 The role of alternation vs. distribution in psycholinguistic studies

Several studies have provided experimental evidence that sounds considered to be variants of the same phoneme are processed differently from sounds considered to be contrastive phonemes (e.g., Beckman & Pierrehumbert 2000; Sumner & Samuel 2005; Kazanina et al. 2006; Ernestus & Baayen 2007), and that speakers find sounds considered to be allophones of a single phoneme harder to discriminate or identify than sounds considered to be separate phonemes (e.g., Lisker & Abramson 1970; Lasky et al. 1975; MacKain et al. 1981; Werker & Lalonde 1988; Lisker 2001). For example, Beckman & Pierrehumbert (2000) asked speakers of English and Korean to identify tokens of Korean [sɨ], [s'ɨ], [ɕi], and [ɕ'i] syllables (where the apostrophe indicates a tense or fortis sound) as containing either of the coronal fricatives *s* and *sh* (using the English orthography). As discussed above, Korean dental fricatives [s/s'] do not occur before the high front vowel [i], and palatal fricatives [ɕ/ɕ'] do not occur before [ɨ]; furthermore, the two sets of fricatives participate in morphological alternations. The results showed that the English listeners successfully identified [s/s'] tokens as *s*, and [ɕ/ɕ'] tokens as *sh*, while the Korean speakers identified the tokens at a chance level. Beckman & Pierrehumbert concluded that the successful identification of *s* and *sh* from the English speakers reflected the phonemic status of the two sounds in English, in which *s* and *sh* may occur in the same environment and signal lexical differences, as in ***s**ee* vs. ***sh**e*. The chance-level rate of correct identification from the Korean speakers reflected the non-phonemic status of the two sounds in Korean, where dental fricatives and palatal fricatives are in complementary distribution and participate in rich and regular morphological alternations.

Along with the experimental evidence suggesting that variants of the same phoneme are harder to discriminate/identify than sounds considered to be contrastive phonemes, there is some evidence that *alternations* may lead speakers to have greater difficulty in discriminating the alternating sounds *even when the sounds would not otherwise be considered allophonic* (Huang 2001; Pierrehumbert 2006b; Ernestus & Baayen 2007). For example, in Mandarin, tone alternations may affect discrimination of tones that are normally contrastive. The difference between Tone 2 (mid-rising, 35) and Tone 3 (low-falling-rising, 214) is contrastive, signaling lexical differences, as shown in the minimal pair *m**á*** 'hemp, T2' and *m**ǎ*** 'horse, T3'. However, in a sequence of two T3 syllables, the contrast is neutralized, with the first T3 becoming T2 (Chao 1968; Shih 1997), as shown in the following examples.

(1) T3→T2 / _ T3 in Mandarin

a. h**ǎ**o	'good'; T3	+	y**ǔ**	'rain'; T3	→	[h**á**o yǔ]	'good rain'; T2 +T3
b. h**á**o	'big'; T2						'big rain'; T2+T3

In (1a), the same morpheme *h**ǎ**o* is realized as T2 *h**á**o* when compounded with another T3 morpheme *y**ǔ***.

Huang (2001) showed that when asked to determine whether two tones were the same or different in an AX paradigm consisting of pairs of single syllables, Mandarin-speaking listeners responded more slowly to pairs containing T2 and T3 than to other pairs of tones. Hume & Johnson (2003) argued that this perceptual confusability results from the fact that the contrast between T2 and T3 is neutralized before T3, even though the tones were not presented in this environment in the experiment. Hume & Johnson thus conclude that the *predictability* of the two tones before T3 reduces perceptual distinctiveness for native listeners.

The discussion in this section and the previous section makes clear that there has not been general agreement on a set of criteria defining phonological relationships. Predictable distribution, the traditional definition of contrast, is not without challenges, particularly since the distribution is not a simple all-or-nothing notion (Hall 2009), as illustrated by the partial predictability/neutralization case of Mandarin T2/T3 mentioned above. And although morphological alternation can reinforce distributional predictability, there is not necessarily a perfect correspondence between distribution and the presence of alternations, as illustrated by the Akan [k]/[tç] case. This dissertation examines the relative contributions of morphological alternation and distribution in sound memberships.

1.3 Languages

To tease apart the relative contributions of *distribution* and *alternation* in motivating speakers to assign sounds to phonological categories, this dissertation compares the behavior of speakers of Mandarin, in which *s* and *sh* are in complementary distribution but do not participate in alternations; English, in which these sounds are not predictable from the phonological environment but show limited alternation; and Korean, in which these sounds are in complementary distribution and regularly alternate. This section presents background on the languages that are investigated in this dissertation.

1.3.1 English

English *s* (/s/) and *sh* (/ʃ/) may occur in the same contexts, giving rise to minimal pairs such as ***s**ea* [**s**i] vs. ***sh**e* [**ʃ**i]. These sounds may alternate optionally at the phonetic level when /s/ is followed by a palatal (*miss* [mɪs] ~ *miss you* [mɪʃju]), and in morphological contexts associated with a small set of derivational suffixes (*oppress* [opɹɛ**s**] ~ *oppression* [opɹɛ**ʃ**ən]; *press* [pɹɛ**s**] ~ *pressure* [pɹɛ**ʃ**ɚ]).[3] However, Johnson & Babel (2010: 129) note that "alternations of this type are infrequent in English and the phonemic contrast between /s/ and /ʃ/ is a very salient aspect of the English phonological system. In English /ʃ/ cannot be derived from [sj]—underlying /Cj/ is only allowed before /u/ in words like *muse*, and /s/ and /ʃ/ contrast in final position where /j/ is phonotactically excluded, as in *lass* [læs], *lash* [læʃ], etc." Thus, English will be considered a case where distributional evidence supports the view that these sounds constitute separate categories, while evidence for grouping them together is weak.[4]

1.3.2 Korean

Korean provides a case in which both distribution and alternation point to the analysis of *s* ([s]) and *sh* ([ɕ]) as members of a single category. The two sounds occur in distinct

---

[3] Zsiga (1995) argues from acoustic and electropalatographic data that post-lexical palatalization is a different process from lexical palatalization.

[4] From Wiktionary (http://en.wiktionary.org/wiki/Appendix:English_suffixes, retrieved on 6/27/2012, 8:56PM), there are 312 derivational suffixes listed, among which, 167 are vowel initial. Only 4 out of the 167 vowel-initial suffixes trigger palatalization (i.e., -*ial*, -*ion*, -*ious*, -*ure*), with half of which providing the pre-palatal contexts (-*ial* and -*ious*).

environments: [ɕ] occurs before the high front vowel [i] and glide [j] (mainly in loanwords), and [s] occurs elsewhere (Sohn 1999; Iverson & Lee 2006; Kim 2009). Table 1-3 provides the Korean consonant inventory with the target fricatives shaded and Table 1-4 illustrates that [ɕ] occurs only before [i] or [j], as in (a)-(f), and [s] occurs elsewhere, as in (g)-(i).

Table 1-3 Korean consonant inventory

	Stop			**Fricative**		**Liquid**	**Nasal**	**Glide**
	Lax	Asp.	Tense	Lax	Tense			
Labial	p	p^h	p’				m	w
Dental	t	t^h	t’	s	s’	l	n	
Palatal	tɕ	$tɕ^h$	tɕ’	ɕ	ɕ’			j
Velar	k	k^h	k’				ŋ	
Glottal				h				

Table 1-4 Complementary distribution of Korean [s] and [ɕ]

a.	[**ɕ**i]	‘poem’
b.	[**ɕ**ikan]	‘time’
c.	[**ɕ**jamp^{h}u]	‘shampoo’
d.	[**ɕ**jap]	‘shop’
e.	[**ɕ**jup^{h}ʌ]	‘super’
f.	[**ɕ**jo]	‘show’
g.	[**s**al]	‘flesh’
h.	[**s**ul]	‘alcohol’
i.	[**s**e]	‘bird’

Korean [s] and [ɕ] also alternate before different vowel suffixes, as shown in Table 1-5. Before the suffix *-e*, indicating locative case, [s] occurs; before the suffix *-i*, indicating nominative case, *the same morpheme* is realized with [ɕ].

Table 1-5 Morphological alternation of [s] *and* [ɕ] *in Korean*

/nas/	[na**s**-e]	'sickle-locative'
	[na**ɕ**-i]	'sickle-nominative'
/kos/	[ko**s**-e]	'place-locative'
	[ko**ɕ**-i]	'place-nominative'
/pus/	[pu**s**-e]	'writing brush-locative'
	[pu**ɕ**-i]	'writing brush-nominative'

Unlike the *s* and *sh* in English, Korean *s* and *sh* occur in predictable environments, and participate in productive morphological alternations. In other words, Korean speakers not only see Clark Kent and Superman in distinct environments (complementary distribution), but also see the same individual enter a phone booth as Clark Kent and leave as Superman.

However, the assumption of the perfect complementary distribution of *s* and *sh* ([ɕ]) in Korean has been questioned, based on the occurrence of words transcribed as [**ɕj**ap] 'shop', [**ɕj**up^{h}ʌ] 'super', and [**ɕj**o] 'show' (c.f., (d)-(f) in Table 1-4). Given the lack of contrast between [ɕ] and [ɕj], such forms might be analyzed as either underlying /sj/ realized as [ɕ] (see the summary of possible analyses of the consonant-glide combination in Suh (2009b: 4), and the references there), or as containing the palatal fricative before back vowels, such as [**ɕ**ap] 'shop', [**ɕ**up^{h}ʌ] 'super', and [**ɕ**o] 'show. In the latter analysis, Korean *s* and *sh* would not be considered to be in perfect complementary distribution since in the context of back vowels, both *s* and *sh* may occur (e.g., [**ɕa**p] 'shop' vs. [**sa**p] 'shovel'). However, even if the two sounds *s* and *sh* are not in full complementary distribution in Korean, their distribution is still restricted, in that *s* never occurs before high-front vowel [i] and *sh* never occurs before non-round mid vowels [ɨ].

1.3.3 Mandarin

In Mandarin, four series of *phonetically similar* sounds—dentals, palatals, retroflexes and velars, as shown in the shaded box in Table 1-6—are *in complementary distribution* (e.g., Chao 1934; Hartman 1944; Cheng 1968; Yip 1996; Duanmu 2007; Wan 2010).

Table 1-6 Mandarin consonant inventory

Labial	p	p^h	f	m
Alveolar	t	t^h	l	n
Dental	ts	ts^h	s	
Palatal	tɕ	tɕʰ		
Velar	k	k^h	x/h	ŋ
Retroflex	tʂ	tʂʰ	ʂ	ʐ

The dental, velar, and retroflex series never occur before high front vowels [i, y] and their corresponding glides [j, ɥ], and palatal sibilants never occur before non-high front vowels or glides, as shown in Table 1-7 (Cheng 1973; Duanmu 2007).

Table 1-7 Complementary distribution of Mandarin fricatives

ɕ tɕ tɕʰ	always before *high-front* vowels [i/y]or glides [j/ɥ] (e.g., [ɕi] 'wash'; [ɕja] 'blind'; [ɕjo] 'rest'; [ɕɥe] 'snow')
x/h k kʰ ʂ tʂ tʂʰ s ts tsʰ	never before [i/y] or [j/ɥ] (e.g., [sa] 'spread'; [so] 'gather')

The distribution of [s] and [ɕ] in Mandarin is similar to the distribution of these sounds in Korean: [ɕ] occurs before high-front vowels/glides, and [s] elsewhere. However, due to Mandarin's lack of affixation and its stringent restrictions on possible syllable structures, the sounds in these series *never display alternations*.[5] Mandarin *s* ([s]) and *sh* ([ɕ]) thus provides a good comparison with the comparable sounds in the other two languages.

However, as in Korean, there has been disagreement concerning the perfect complementary distribution of Mandarin [s] and [ɕ]. Li (2008) states that in the context of the vowels /a/ and /o/, [s] and [ɕ] are robustly contrastive because there is no obvious glide present after [ɕ] ([sa] 'spread' vs. [ɕa] 'blind'; [so] 'gather' vs. [ɕo] 'rest'; c.f., Table 1-7) (Li 2008: 17). On the other hand, Duanmu (2007) argues on the basis of the distribution of glides that the reason why there is no obvious glide present after [ɕ] is that Mandarin onsets have only a single slot, which a consonant and glide must share, and [ɕ] is actually a surface realization of the consonant-glide combination /sj/ (/sa/→[sa] 'spread' vs. /**sj**a/→[**ɕ**a] 'blind'; /so/→[so] 'gather' vs. /**sj**o/→[**ɕ**o] 'rest'). Duanmu's argument, along with others (e.g., Chao 1934; Hartman 1944; Cheng 1968; Lin 1989; Chiang 1992; Wu 1994; Yip 1996), suggests that Mandarin [s] and [ɕ] are in complementary distribution. However, whether *s* and *sh* in Korean and Mandarin are in perfect complementary distribution, or whether they overlap in some but not all contexts, the distributions of the two sounds in Korean and Mandarin are restricted, and different from that of these sounds in English, in which these sounds can occur in the same context. Furthermore,

[5] Only dental nasal /n/ and velar nasal /ŋ/ can occur in coda position, and codas do not re-syllabify to a following onsetless syllable. The fricatives in question do not occur in coda position.

unlike Korean, these sounds in Mandarin never display alternations. Mandarin [s] and [ɕ] thus provide a good comparison with the comparable sounds in the other two languages.

The different types of relationships between the coronal fricatives *s* and *sh* (English [ʃ] and Korean/Mandarin [ɕ]) in these three languages are summarized in Table 1-8. The parentheses around the English checkmark indicate that the morphological alternations of *s* and *sh* are limited in English (see section 1.3.1).

Table 1-8 Languages

	Mandarin	Korean	English
Distribution	√	√	
Alternation		√	(√)

The differences in the application of the factors used to define phonological relationships of these three languages will help us determine the relative contributions of *predictability* and *alternation* in leading speakers to group sounds as members of the same phoneme category. Since Korean provides a case in which both distribution and alternation point to the analysis of *s* and *sh* as members of a single category, these sounds in Korean will be taken as a baseline. I consider the following somewhat simplified hypotheses to guide the discussion, *Distribution Alone Hypothesis*, *Alternation Alone Hypothesis*, and *Distribution Plus Alternation Hypothesis*. If distributional predictability is a sufficient condition for causing speakers to group sounds into a single phonological category (i.e., the case of *s/sh* in Mandarin)—in other words, if distribution alone is as strong as distribution plus alternation (i.e., the case of *s/sh* in Korean)—we expect the results from the Mandarin group to be similar to those from the Korean speakers, and to be different from those of the English speakers (*Distribution Alone Hypothesis*). If alternation is a sufficient condition to cause speakers to group sounds together (i.e., the case of *s/sh* in English)—in other words, if alternation alone is as strong as distribution plus alternation—then English and Korean speakers should pattern similarly (though probably in a way that reflects the much weaker evidence from alternations in English) (*Alternation Alone Hypothesis*). If distribution, as well as alternation, is necessary to cause speakers to group sounds as a single category—in other words, having one or the other is not sufficient—then we expect the results from the Mandarin group and from the English group to pattern similarly, and to be different from the Korean group (*Distribution Plus Alternation Hypothesis*). In the dissertation, I will not pursue the *Alternation Alone Hypothesis*, based on previous findings showing that English speakers' perception reflects the contrastive status of *s* and *sh*. Johnson & Babel (2010), in a similarity rating experiment, showed that English participants rated *s* and *sh* as more different than did Dutch participants, in whose language *s* and *sh* are considered to be variants of the same phoneme category. Johnson & Babel concluded that the different rating patterns reflect the different phonological status of the two sounds in English and in Dutch: *s/sh* are contrastive in English but are allophonic variants of the same phoneme in Dutch (see more on this experiment in Chapter 2). Furthermore, as mentioned in section 1.2, English participants, in an identification task (Beckman & Pierrehumbert 2000), showed categorical perception of *s* and *sh* while Korean participants showed only chance level perception, corresponding to the non-phonemic status of *s* and *sh* in their native language. These previous findings suggest that the limited morphological alternations of *s*/*sh* in English are not sufficient to draw a comparison with the two sounds in Korean, in which *s/sh* participate in rich and regular morphological alternations, and that English listeners' perception reflects the contrastive status of *s* and *sh*.

The predictions of the two hypotheses are shown in Table 1-9 with the predicted patterns among the three languages boxed.

Table 1-9 Predictions

Distribution Alone Hypothesis

	Mandarin	Korean	English
Distribution	√	√	
Alternation		√	(√)

Distribution Plus Alternation Hypothesis

	Mandarin	Korean	English
Distribution	√	√	
Alternation		√	(√)

The experiments presented in the following chapters yield mixed but interesting results, suggesting that there is no simple answer to the question of whether two sounds are members of the same category. Rather, the relationship between two sounds is determined by multiple factors and is not always absolute and categorical (Goldsmith 1995; Hall 2009).

The implications of the findings in the dissertation are threefold. First, the findings have implications for the criteria used to determine category membership cross linguistically. Second, the findings provide psycholinguistic evidence relevant to the long-standing debate on the analysis of Mandarin palatals. Third, the findings shed light on the assumptions and definitions of different phonological models. The implications for different phonological theories will be discussed in Chapter 5.

The rest of the dissertation is structured as follows. Chapter 2 presents the experiment on similarity ratings of *s*/*sh* by the three language groups, showing that overall, Mandarin and

English speakers' ratings of *s* and *sh* were more different from those of speakers of Korean. The results suggest that Mandarin speakers, who have access only to *distributional* evidence, are less likely to treat *s/sh* as members of a single category than Korean speakers, who have access to *distributional* evidence as well as *morphological alternation*, supporting the *Distribution Plus Alternation Hypothesis*. Furthermore, the results also showed that the similarity judgments from the speakers of all three languages varied in different vowel contexts, suggesting that phonological relationships are *gradient* rather than categorical and depend on *multiple* factors. Chapter 3 presents the experiment on discrimination of *s*/*sh* by English, Korean, and Mandarin speakers. The accuracy results of this experiment did not yield a difference in the ability to discriminate the two sounds according to language background. These results seem to contradict the results of the first study, in which English and Mandarin speakers patterned together in contrast to Korean speakers. Furthermore, the response time results showed that the English speakers patterned with the Korean speakers in that, overall, they took less time than did the Mandarin speakers in their discrimination of *s/sh*. In Chapter 3 I discuss possible explanations for the discrepancy between the results of the similarity rating and discrimination experiments. Chapter 4 presents the experiment on the semantic priming of *s* and *sh* by English, Korean, and Mandarin speakers. The results showed that English *s* and *sh* did not exhibit a priming relation while Mandarin and Korean *s* and *sh* did. The results again seem to contradict the *Distribution Plus Alternation Hypothesis*, which predicts that the Mandarin speakers should pattern with the English speakers and deviate from the Korean speakers. I will discuss possible explanations of these results in Chapter 4. Finally, Chapter 5 provides the implications of the findings and concludes the dissertation.

Chapter 2 Similarity Ratings

This chapter presents the results of a similarity rating experiment designed to investigate how English, Korean, and Mandarin speakers perceive the relative similarity of the target fricatives, *s* and *sh*. The experiment was designed based on research showing that sounds that are allophonic variants of the same phoneme in the participant's native language are perceived as more similar than sounds that belong to different phoneme categories (Harnsberger 2001; Boomershine et al. 2008; Babel & Johnson 2010).

In English, *s* and *sh* occur in the same context and substituting one for the other may signal a difference in meaning. In Korean *s* and *sh* are in complementary distribution and participate in regular and productive morphological alternations. In Mandarin, *s* and *sh* are in complementary distribution but do not participate in allomorphic alternations. In a distributionally based approach in which sounds in complementary distribution are considered variants of the same category, Mandarin *s/sh* should have the same status as Korean *s/sh*. From this point of view, we predict that Korean and Mandarin speakers should rate *s* and *sh* as more similar than English speakers do (*Distribution Alone Hypothesis*). However, if distribution alone is not assumed to force learners to map sounds in complementary distribution onto the same underlying category, we predict that Mandarin speakers, just like English speakers, should rate *s* and *sh* as more different than Korean speakers do (*Distribution Plus Alternation Hypothesis*).

The results of the experiment showed that Mandarin speakers' ratings of *s* and *sh* differed from those of Korean speakers, and patterned overall with those of the English speakers. These results are consistent with the *Distribution Plus Alternation Hypothesis*, suggesting that the Mandarin speakers, who have access only to *distributional* evidence, are less likely to treat *s/sh* as members of a single category than the Korean speakers, who are exposed to evidence from both *distribution* and *morphological alternation*. However, an unexpected effect emerged: similarity judgments from all three language groups varied in different vowel contexts, supporting the view that sound category membership is not simply all-or-nothing (Goldsmith 1995; Hall 2009).

The structure of this chapter is as follows. Section 2.1 provides a review of the previous literature on similarity ratings. Section 2.2 describes the methodology of the experiment, followed by results in section 2.3. Section 2.4 provides a summary.

2.1 Introduction

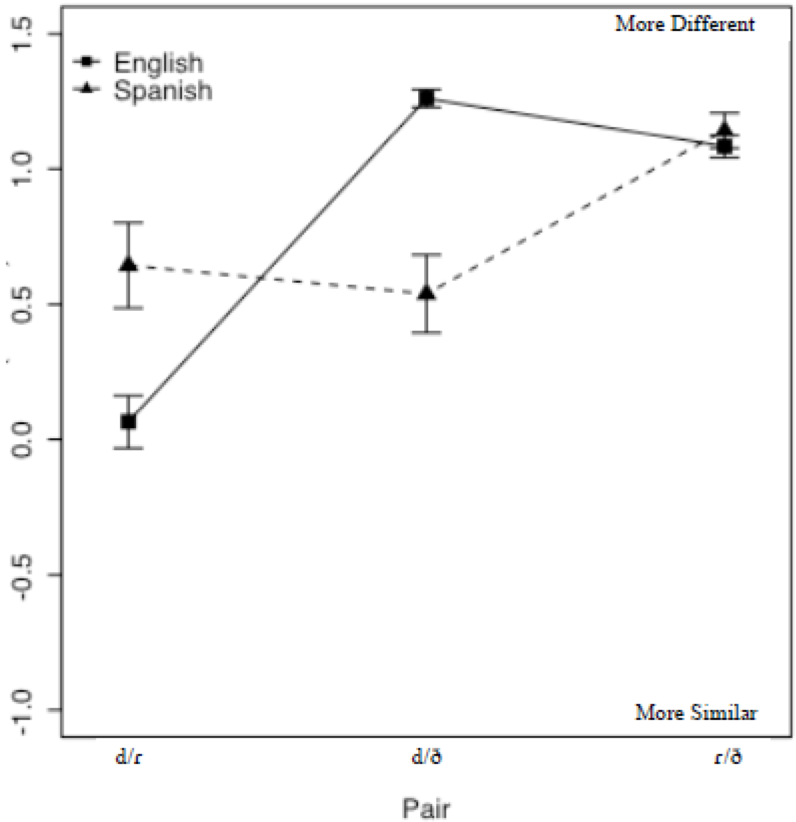
In similarity rating tasks, speakers tend to rate sounds in an allophonic relationship as more similar than separate phonemes (Harnsberger 2001; Boomershine et al. 2008; Babel & Johnson 2010; Johnson & Babel 2010). For example, Boomershine et al. (2008) tested native English and Spanish speakers' similarity judgments of [ð], [d], and [ɾ] in different vowel contexts using an AX paradigm (e.g., [ada]-[aɾa], [idi]-[iði]). [ð] and [d] are contrastive phonemes in English (e.g., *they* [**ð**eɪ] vs. *day* [**d**eɪ]) but are allophonic variants in Spanish, due to a process whereby intervocalic voiced stops are spirantized following a continuant (e.g., [d]*onde* ‘where’ but *de* [ð]*onde* ‘from where’). In contrast, [d] and [ɾ] are contrastive in Spanish (e.g., [kaða] ‘each’, [kaɾa] ‘face’) but are allophonic variants in American English, due to a process whereby [d] (and [t]) become a tap intervocalically preceding an unstressed vowel (e.g., *ride* [raɪ**d**], but *rider* [raɪ**ɾ**ɚ]). The phonological relationships of the three sounds are shown in Table 2-1.

Table 2-1 Phonological grouping of [ð], [d], and [ɾ] in English and Spanish

English	[ð]	[d], [ɾ]
Spanish	[ð], [d]	[ɾ]

Boomershine et al. (2008) asked participants to rate the similarity of a pair of sounds taken from the VCV sequences [ada], [aɾa], [aða], [idi], [iɾi], [iði], [udu], [uɾu], and [uðu]. The vowel context was the same for every pair so that the only difference in each pair was the consonant. Participants rated the pairs on a scale of 1-5, where 1 indicated ‘very similar’ and 5 indicated ‘very different’. The results showed a clear native language effect, as shown in Figure 2-1. The *x* axis represents the fricative pairs, [d/ɾ], [d/ð], and [ɾ/ð]. The *y* axis represents the normalized similarity rating scores (*z*-scores), with scores above zero indicating ‘more different’ and scores below zero indicating ‘more similar.’ When the participants judged the [ɾ/ð] pair (on the right of Figure 2-1), two sounds that are contrastive in both English and Spanish, the rating scores of the two language groups converged. However, for the other two pairs, the English speakers rated [d] and [ɾ] as most similar (left), while the Spanish speakers rated [ð] and [d] as most similar (center), patterning with the phonological relationships of the two sounds in English and Spanish.

Figure 2-1 Similarity rating results of [d], [ð], and [ɾ] in English and Spanish (Boomershine et al. 2008)



Johnson & Babel (2010) tested native English and Dutch speakers' similarity judgments of [s] and [ʃ] (along with other fricatives) using the same methodology as Boomershine et al. (2008). Dutch [s] and [ʃ] participate in morphological alternations (e.g., *poes* [s] 'cat'~ *poesje* [ʃ] 'kitten', and *tas* [s] 'bag'~ *tasje* [ʃ] 'small bag'), and they also alternate in connected speech (*wa*[s j]*e* ~ *wa*[ʃ]*e* 'were you' and *ze*[s j]*anuari* ~ *ze*[ʃ]*anuari* 'January the 6th') (Gussenhoven 1999). Though [ʃ] exists in borrowed words (e.g., *chef* [ʃ] 'chef, boss'; *sjaal* [ʃ] 'shawl'), it is argued to derive from an allophonic rule that palatalizes /s/ before /j/, and thus [ʃ] is analyzed as a variant of /s/ before the high-front glide [j] in Dutch (Booij 1999).[6] English [s] and [ʃ], though they sometimes alternate on a phonetic level (*miss* [mɪ**s**] ~ *miss you* [mɪ**ʃ**jiu]) or through limited morphophonological alternations (*oppress* [opɹɛ**s**] ~ *oppression* [opɹɛ**ʃ**ən]; *press* [pɹɛ**s**] ~ *pressure* [pɹɛ**ʃ**ɚ]), one cannot reliably predict the occurrence of *s* and *sh* from context, and there are a large number of minimal pairs differing only in these sounds (e.g., ***s**ee* vs. ***sh**e*, and ***s**ue* vs. ***sh**oe*). The fricative phonemic inventories of Dutch and English are listed in Table 2-2 with the target sounds shaded.

Table 2-2 Voiceless fricative phonemic inventories of Dutch and English (Babel & Johnson 2010)

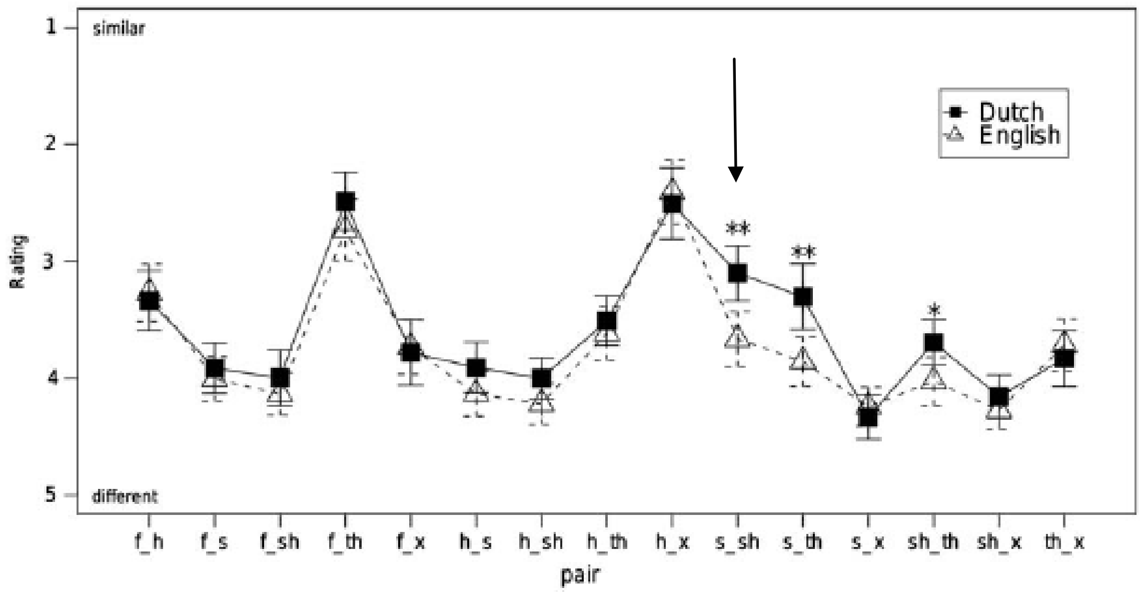
	Labiodental	Dental	Alveolar	Post-alveolar	Velar	Glottal
Dutch	f		s		x	h
English	f	θ	s	ʃ		h

As shown in Figure 2-2, Johnson & Babel also found a native language effect. The *x* axis shows the different fricative pairs and the *y* axis shows the similarity rating scores from 1 to 5, 1 being very similar, and 5 being very different. The perceived difference between [s] and [ʃ] (the pair indicated with an arrow) for their Dutch listeners was significantly smaller (higher similarity

[6] The allophonic rule of palatalization also involves coronal obstruents and nasals /s, z, t, n/ before /j/ (Booij 1999).

ratings) than the phonetic difference reported by English listeners (lower similarity ratings), while the rating scores for other fricative pairs were similar for the two language groups.[7]

Figure 2-2 Similarity rating results of voiceless fricatives in Dutch and English (Babel & Johnson 2010)



Note that the ‘allophonic’ cases presented here (i.e., English [d/ɾ], Spanish [d/ð], Dutch [s/ʃ]) are the ones that are both distributionally predictable and morphologically alternating. Thus, these prior studies do not tease apart these two factors. The experiment in this dissertation tested cases where both distribution and alternation exist side-by-side (*s*/*sh* in Korean), and extended it by looking at language instances where each factor occurs independently (predictable distribution of *s*/*sh* in Mandarin and alternation of *s*/*sh* in English).

[7] For the explanation of other non-converging fricative pairs (i.e., [s]-[θ] and [ʃ]-[θ]), see the discussion in Johnson & Babel (2010).

Based on the findings using similarity rating tasks, if distribution alone defines the phonological relationship of the two sounds, we expect the Mandarin listeners' ratings to be similar to those of Korean listeners and different from those of English listeners (*Distribution Alone Hypothesis*). If distribution and alternation are both necessary in grouping sounds as variants of the same category, we expect the ratings of Mandarin listeners to be different from those of Korean listeners (*Distribution Plus Alternation Hypothesis*). The predictions are summarized in Table 2-3 with the predicted patterns boxed.

Table 2-3 Predictions of similarity ratings

Distribution Alone Hypothesis: For Mandarin and Korean speakers, *s* and *sh* are single category.

Predictions of similarity rating results: Reduced perceptual distance between *s* and *sh* in Mandarin and Korean.

	Mandarin	**Korean**	**English**
Distribution	√	√	
Alternation		√	(√)

Distribution Plus Alternation Hypothesis: For Korean speakers, but not Mandarin speakers, *s* and *sh* are single category.

Predictions of similarity results: Reduced perceptual distance between *s* and *sh* only in Korean.

	Mandarin	**Korean**	**English**
Distribution	√	√	
Alternation		√	(√)

2.2 Methodology

The goal of this set of experiments was to compare how English, Korean, and Mandarin listeners rate the target sounds, *s* and *sh*.

2.2.1 Participants

20 participants from three language groups were recruited for this set of experiments. Participants in the English group (11 male, 9 female, aged 18-22), all monolingual speakers of English, and in the Korean group (6 male, 14 female, aged 18-38), all native speakers of Korean from South Korea, were recruited at Stony Brook University, and received course credit or payment for their participation. To estimate possible influence from English, Korean participants, who had all received up to a high school education in South Korea before coming to Stony Brook, were asked to rate their English ability; the average rating was 4.65 on a 7-point scale (see Appendix A for an example questionnaire). Participants in the Mandarin group (4 male, and 16 female, aged 20-22) were all native speakers of Taiwanese Mandarin, and were recruited at National Chiao Tung University in Taiwan for course credit or payment. Their average self-rating of English ability was 4.6 on a 7-point scale. None reported any hearing deficiencies.

2.2.2 Designs and materials

The materials contained the target fricatives [s, ɕ, ʃ] along with two other fricatives [f, h] as controls, embedded in three vowel contexts [a_a], [i_i], and [u_u]. Materials consisted of two tokens of each of the following VCV sequences: [asa][aɕa][aʃa][afa][aha], [isi][iɕi][iʃi][ifi][ihi], or [usu][uɕu][uʃu][ufu][uhu]. Note that the fricative [ɕ] does not exist in the English consonant inventory, [f] does not exist in Korean, and [ʃ] does not exist in both Korean and Mandarin (c.f., section 1.3). Also note that some of the stimuli contained illicit sequences according to the phonotactics of individual languages: *[si] and *[ɕu] in Korean/Mandarin; *[fi] and *[hi] in Mandarin. The tokens were produced by a trained male phonetician whose native language is Mandarin. The Mandarin native speaker was chosen to record the stimuli because he was able to produce the Korean/Mandarin alveo-palatal fricative [ɕ] and the English [ʃ] from extensive English exposure, and the combinations of these sounds in different vowel contexts from professional training. The speaker recorded multiple examples of the stimuli with high tone on both syllables. One instance of each VCV was selected as a test item so that the tokens were approximately matched on pitch and duration. Table 2-4 shows the average pitch of the first and second vowel of the selected stimuli (V1 mean across vowels: 115.87 Hz, standard deviation: 2.33 Hz; V2 mean across: 116.2 Hz, standard deviation: 2.01 Hz), and Table 2-5 shows the vowel and fricative durations of the selected stimuli (total duration mean: 726.4 ms, standard deviation: 32.28 ms). In order to control the intensity across tokens, the average intensity of each token was scaled to 65 dB, the rough average of the intensity of all the tokens, using Praat software (Boersma 2001).

Table 2-4 Pitch in Hz of the first and second vowel

	V1	V2
aça	118	117
afa	120	118
aha	119	117
asa	114	116
aʃa	114	116
içi	113	115
ifi	114	114
ihi	113	114
iʃi	114	114
isi	114	112
uçu	118	118
ufu	118	119
uhu	116	118
uʃu	116	118
usu	117	117

The design followed closely that of Boomershine et al. (2008) and Johnson & Babel (2010). This set of experiments is a three-factorial design with one between-subject factor (Language), and two within-subject factors (Fricative Pair, Vowel Context), as shown in Table 2-6.

Table 2-5 Durations in ms of the first vowel, the fricative, second vowel, and the total duration of the stimulus

	V1	Fric	V2	Total
aça	225	198	330	753
afa	301	142	333	775
aha	277	128	313	717
asa	266	162	337	765
aʃa	232	168	294	694
içi	255	201	320	776
ifi	262	152	329	743
ihi	278	137	305	720
iʃi	213	216	305	734
isi	243	182	309	734
uçu	194	213	315	722
ufu	226	169	291	685
uhu	231	152	288	671
uʃu	222	227	263	712
usu	203	196	297	695

Table 2-6 Similarity rating design

Between-subject factor	Language	→English, Korean, Mandarin
Within-subject factor	Fricative Pair	→[s-ç], [s-ʃ], [s-f], [s-h] [ç-ʃ], [ç-f], [ç-h] [ʃ-f], [ʃ-h] [f-h]
	Vowel Context	→[a_a], [i_i], [u_u]
Dependent variable	Rating score	→1(similar)-5(different)

The pairing of the five fricatives, setting order aside, gives 5^2= 25 possible pairs, including 5 pairs in which both members were the same (5^2- 5 same pairs= 20 different pairs). Each different pair was presented once, while the same pairs were presented twice to balance the number of same and different pairs, yielding 30 trials (20 different pairs + 5 x 2 same pairs= 30) per vowel context (30 x 3 vowel contexts= 90). Listeners heard each of the AX trials (90 trials) three times in the 3 blocks (90 x 3 blocks = 270) with an inter-stimuli interval (ISI) of 1000 ms between A and X. Participants had a maximum of 5000 ms before the next trial started if they did not respond to a given trial.

2.2.3 Procedure

Participants were presented with written instructions on the computer screen in their native language saying that they would hear a pair of sounds and be asked to rate how similar those sounds were on a scale of 1 to 5, where 1 was 'very similar' and 5 was 'very different.' The participants took part in the experiments individually, or in groups of up to four people in separate booths, using a computer that was connected to a keyboard with 5 keys labeled from 1 to 5. The pairs were presented in different random orders for each participant, using E-Prime software (v2.0; Psychological Software Tools, Pittsburgh, PA). All stimuli were presented binaurally over headphones at a comfortable listening level. The participants completed a 9-trial practice randomly chosen from the test trials, and had the opportunity to ask questions before proceeding to the experiment. The experiment lasted approximately 20 minutes.

2.3 Results

We expect more different ratings between *s* and *sh* for the English and more similar ratings between *s* and *sh* for the Korean listeners (Beckman & Pierrehumbert 2000). Of particular interest here are the Mandarin listeners' ratings. If distribution alone defines the phonological relationship of the two sounds, we expect the Mandarin listeners' ratings to be similar to those of the Korean listeners (*Distribution Alone Hypothesis*). If alternation contributes in defining phonological relationships, we expect the ratings of the Mandarin listeners to be more different from those of the Korean listeners (*Distribution Plus Alternation Hypothesis*).

The rating scores for each participant were normalized into *z*-scores (the difference between the individual score and the mean divided by standard deviation) to compensate for differences in using the 5-point scale (Boomershine et al. 2008). The standardized scores were centered around zero, with scores above zero indicating ‘more different’ and scores below zero indicating ‘more similar.’ The normalized results are shown in Table 2-7, and illustrated in Figure 2-3.[8] The *x* axis represents the different fricative pairs and the *y* axis represents the normalized *z*-scores.

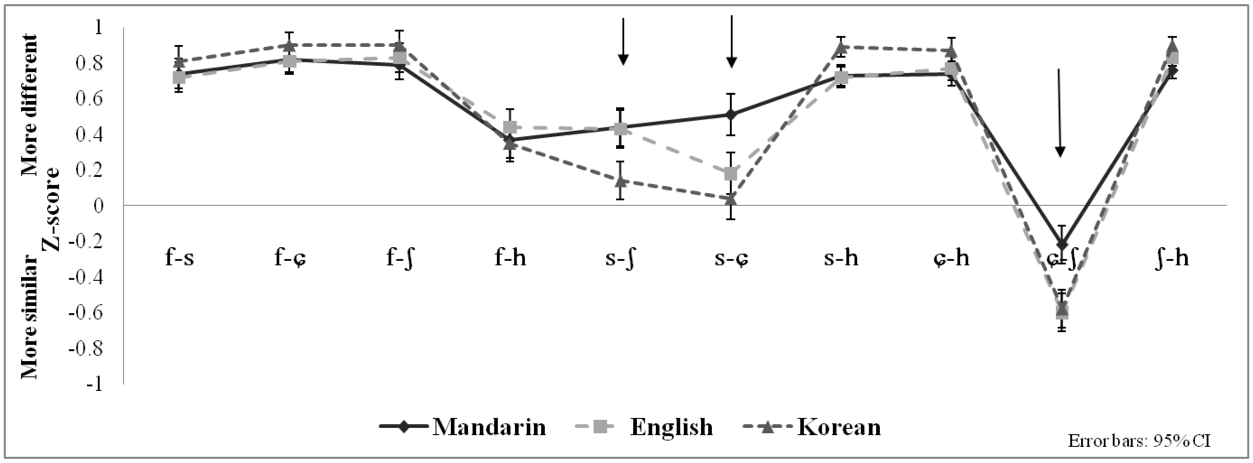
Table 2-7 Similarity rating normalized results

Pair / Language	f-s	f-ɕ	f-ʃ	f-h	s-ʃ	s-ɕ	s-h	ɕ-h	ɕ-ʃ	ʃ-h
Mandarin	0.74	0.82	0.79	0.37	0.44	0.51	0.73	0.74	-0.22	0.76
English	0.72	0.81	0.83	0.44	0.43	0.18	0.72	0.77	-0.60	0.83
Korean	0.81	0.90	0.90	0.35	0.14	0.04	0.89	0.87	-0.58	0.90

From Figure 2-3, we can see that except for the target pairs indicated by the arrows, ratings from the three language groups were very similar. The differences lie in the target pairs [s-ʃ], [s-ɕ], and [ɕ-ʃ]. I will discuss the effect of language in section 2.3.1, the effect of vowel context in section 2.3.2, and the overall results in section 2.3.3.

[8] The assumption of normality was met with the *z*-score transformed results. See Appendix E.1.

Figure 2-3　　Similarity rating normalized results



2.3.1　Effect of language

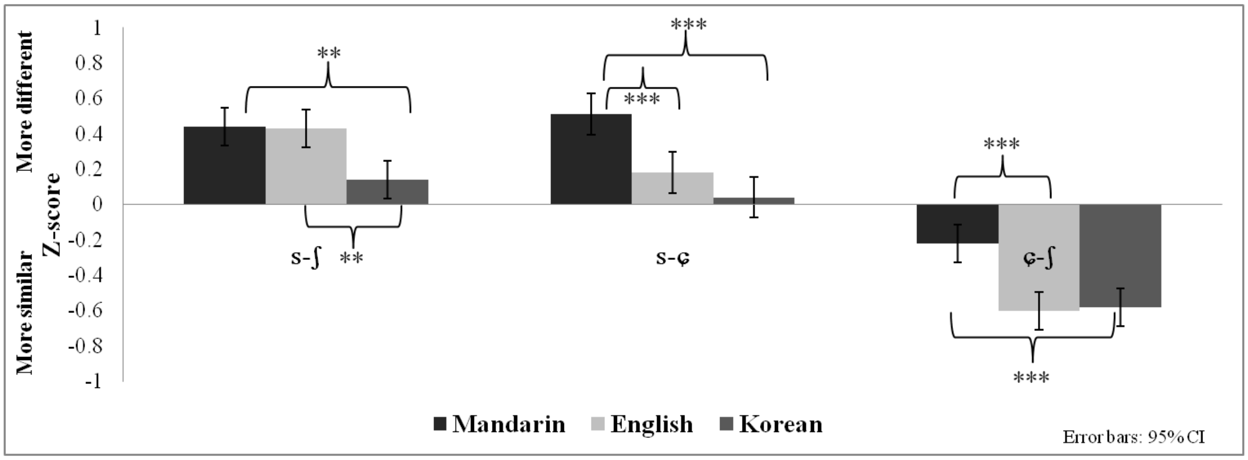
A repeated-measures analysis of variance (ANOVA) (**Language** [Mandarin, English, Korean] **x Fricative Pair** [s, ɕ, ʃ, f, h]) was performed to interpret the results. The analysis showed that there was a main effect of Language (*F*(2,57)=7.962, *p*=.001), and of Fricative Pair (*F*(9,513)=273,419, *p*<.001). In other words, the ratings differed for different language groups, as well as for different fricative pairs. Most importantly, there was a significant Fricative Pair by Language interaction (*F*(18,513)=8.647, *p*<.001), meaning that the ratings for pairs of fricatives were statistically different depending on the native language of the participants. The statistical results are summarized in Table 2-8 (*: *p*<.05; **: *p*<.01; ***: *p*<.001).

Table 2-8　　Summary of similarity rating results

**Language
***Fricative Pair
***Language x Fricative Pair

Of interest here are the ratings of the target pairs, [s-ʃ], [s-ɕ], and [ɕ-ʃ], as shown in Figure 2-4. The *x* axis represents the three target fricative pairs and *y* axis represents the normalized *z*-scores.

Figure 2-4 S/sh similarity rating results



The [s-ʃ] and [s-ɕ] pairs (the two *s-sh* pairs) were rated as more similar by the Korean group than by the English and Mandarin groups. Subsequent analyses showed that the factor Language was significant for the [s-ʃ] pair ($F(2,57)=10.243$, $p<.001$). Post-hoc tests showed that the significance came from Mandarin vs. Korean, and English vs. Korean (both $p<.01$).[9] The ratings from Mandarin vs. English were not significantly different ($p=.991$). That is, the Mandarin and English groups patterned the same for the [s-ʃ] pair, while the Korean group rated these sounds as significantly more similar than the other two groups.

[9] Tukey procedure was used throughout the dissertation for post-hoc tests to control the family-wise error rate over the entire set of pairwise comparisons.

The factor Language yielded a significant effect in the [s-ɕ] pair as well ($F(2,57)=17.510$, $p<.001$). The significance came from Mandarin vs. English and Mandarin vs. Korean from post-hoc tests (both $p<.001$). The standardized rating scores by the English group were higher than those by the Korean group (meaning [s] and [ɕ] were more different for English listeners than Korean listeners), though the difference was not significant ($p=.214$). For the [s-ɕ] pair, though the Korean and English groups patterned similarly, this pattern was induced by a certain vowel context (i.e., [i_i] context; see discussion in sections 2.3.2 and 2.3.3).

As for the ratings for the [ɕ-ʃ] pair, two kinds of *sh* ([ɕ] in Korean and Mandarin, [ʃ] in English), we can see from Figure 2-4 that listeners from all three languages rated them as very similar (all below 0), though Mandarin listeners' ratings were higher overall (i.e., more different). Subsequent analyses showed that the factor Language was significant in the [ɕ-ʃ] pair ($F(2,57)=15,859$, $p<.001$). Post-hoc tests showed that the significance came from the Mandarin vs. English groups and the Mandarin vs. Korean groups (both $p<.001$). There was no significant difference between the English vs. Korean groups ($p=.967$). A possible explanation for why the Mandarin speakers rated the [ɕ-ʃ] pair as less similar than the other two groups will be provided in the next section (2.3.2).

The statistical results for the target fricative pairs are summarized in Table 2-9.

Table 2-9 Summary of similarity ratings on s/sh pairs

***Simple effect of Language in [s-ʃ]	**Mandarin & Korean **English & Korean Mandarin & English (*p*=.991)
***Simple effect of Language in [s-ɕ]	***Mandarin & English ***Mandarin & Korean English & Korean (*p*=.214)
***Simple effect of Language in [ɕ-ʃ]	***Mandarin & English ***Mandarin & Korean English & Korean (*p*=.967)

To summarize the overall results, the [s-ʃ] pair was rated as more different by the Mandarin and English groups than by the Korean group; the [s-ɕ] pair was rated as more different by the Mandarin group than by the English and Korean groups. The [ɕ-ʃ] pair was rated as more similar by the English and Korean groups than by the Mandarin group.

2.3.2 Effect of vowel context

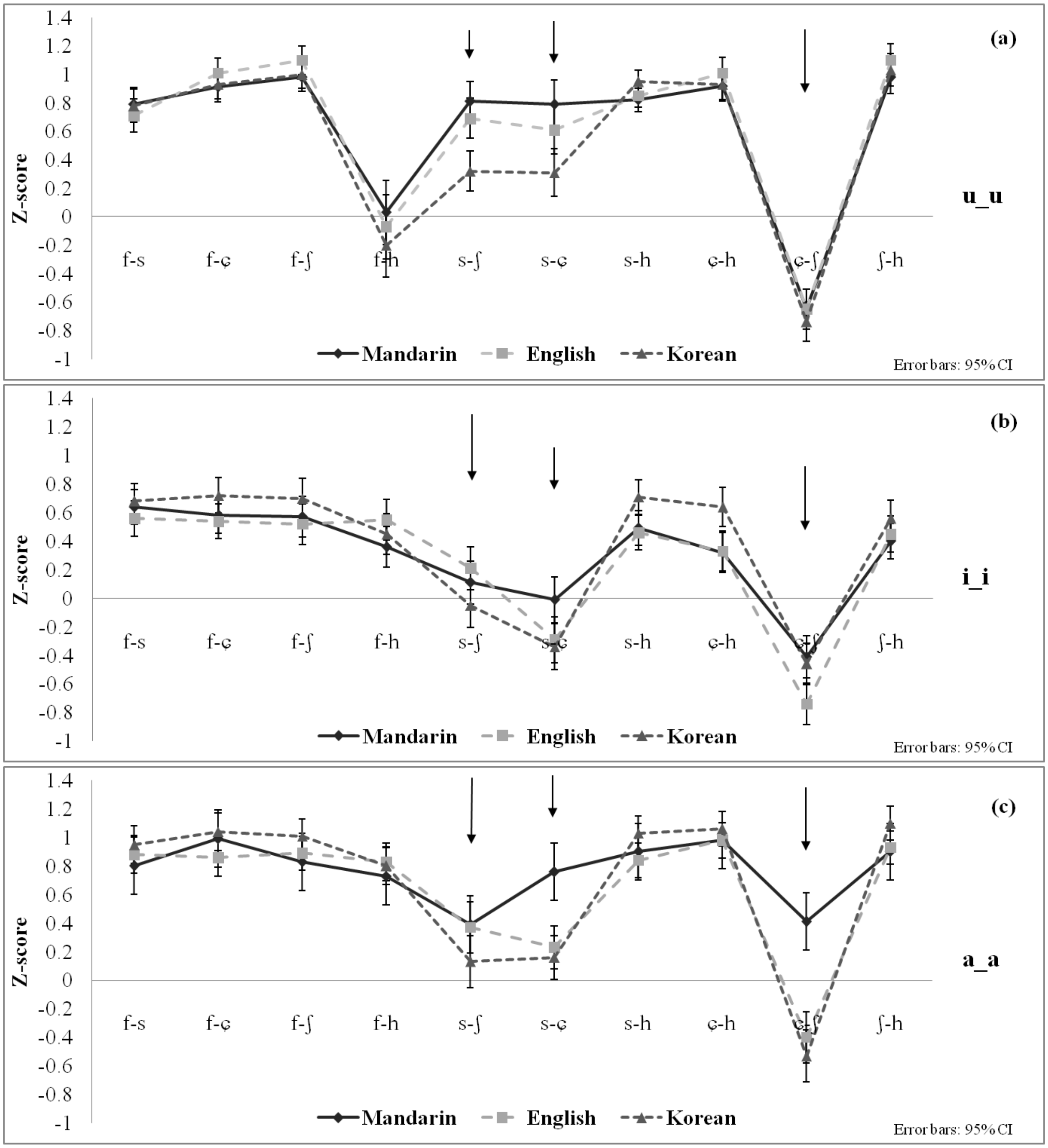
Another repeated-measures ANOVA including Vowel Context as a variable (**Language**: Mandarin, English, Korean **x Pair**: [s, ɕ, ʃ, f, h] **x Vowel Context** [a_a], [i_i], [u_u]) yielded some unexpected results, as shown in Table 2-10 and Figure 2-5.

Table 2-10 Similarity rating standardized results by Vowel Context

Language \ **Pair**		**f-s**	**f-ɕ**	**f-ʃ**	**f-h**	**s-ʃ**	**s-ɕ**	**s-h**	**ɕ-h**	**ɕ-ʃ**	**ʃ-h**
Mandarin	a_a	.80	.99	.83	.73	.39	.76	.90	.98	.41	.90
	i_i	.64	.58	.57	.36	.11	-.01	.49	.32	-.41	.40
	u_u	.79	.91	.98	.03	.81	.79	.82	.92	-.65	.98
English	a_a	.88	.86	.89	.83	.37	.23	.84	.98	-.40	.93
	i_i	.56	.54	.52	.55	.21	-.29	.46	.33	-.74	.45
	u_u	.71	1.01	1.10	-.07	.69	.61	.85	1.01	-.65	1.10
Korean	a_a	.95	1.04	1.01	.80	.13	.16	1.03	1.06	-.53	1.10
	i_i	.68	.72	.70	.45	-.05	-.34	.71	.64	-.46	.56
	u_u	.78	.93	1.00	-.20	.32	.31	.95	.93	-.74	1.03

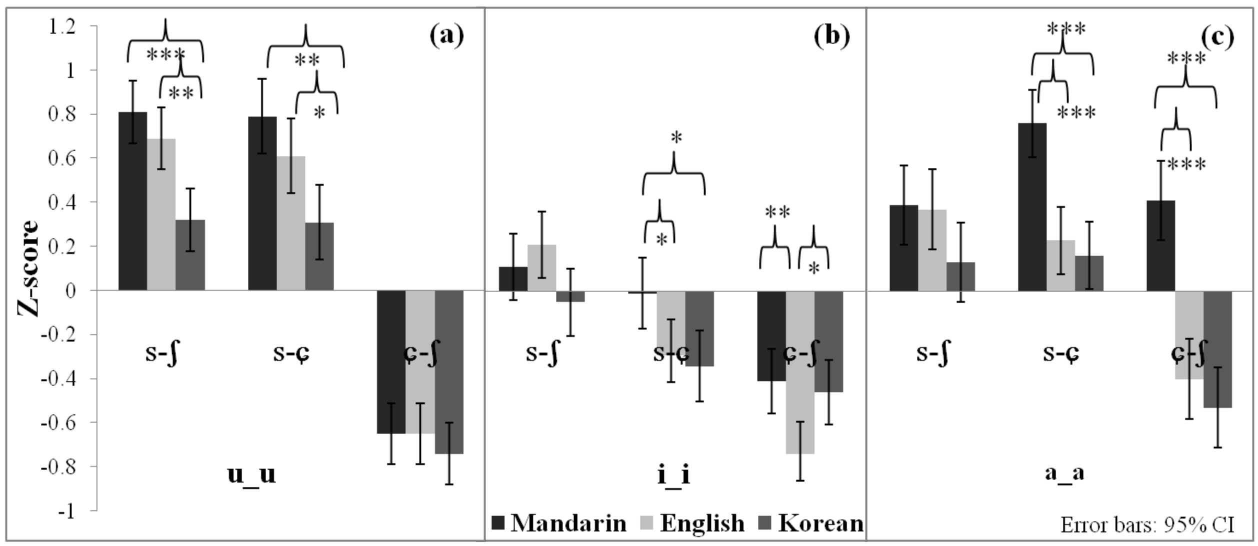
The *x* axis in Figure 2-5 represents the different fricative pairs and the *y* axis represents the normalized *z*-scores. From Figure 2-5 ((a): [u_u]; (b): [i_i]; (c): [a_a]), we can again see that the similarity ratings differed mainly in the target pairs (indicated by the arrows).

Figure 2-5 Similarity rating standardized results by Vowel Context



Of interest here are the ratings of the target pairs, [s-ʃ], [s-ɕ], and [ɕ-ʃ], as shown in Figure 2-6 ((a): [u_u]; (b): [i_i]; (c): [a_a]).

Figure 2-6 S/sh similarity rating results by Vowel Context



For the Korean group, a trend toward rating all the target pairs as more similar than the other two groups was observed, except for the [ɕ-ʃ] pair. Subsequent analyses showed that in the [u_u] context (Figure 2-6 (a)), the Mandarin group patterned with the English group in that the Mandarin and English listeners rated the target pairs as more different than the Korean listeners. The factor Language was significant in the [s-ʃ] pair ($F(2,60)=13.124$, $p<.001$), and post-hoc tests showed that the significance came from the difference of the Mandarin vs. Korean groups ($p<.001$), and of the English vs. Korean groups ($p<.01$). No difference was found between the Mandarin vs. English groups ($p=.459$). A significant Language effect was found in the [s-ɕ] pair as well ($F(2,60)=17.510$, $p<.001$). The significance came from the difference of the Mandarin vs. Korean groups ($p<.01$), and of the English vs. Korean groups ($p<.05$). No effect was found for the [ɕ-ʃ] pair ($F(2,60)=.619$, $p=.542$). The results for the [u_u] context showed that the Mandarin and English listeners treated [s] and [ʃ/ɕ] as more different from each other than did Korean listeners.

In the [i_i] context (Figure 2-6 (b)), the Korean listeners again rated the target pairs as more similar than the other two groups, except for the [ɕ-ʃ] pair: the English listeners showed the most similar ratings for the [ɕ-ʃ] pair. No statistical difference was found for the [s-ʃ] pair, though a trend of more similar ratings from the Korean group was present ($F(2,60)=2.956$, $p=.06$). The factor Language was significant in the [s-ɕ] pair ($F(2,60)=4.966$, $p<.05$), and the significance came from the difference of the Mandarin vs. English groups, and of the Mandarin vs. Korean groups (both $p<.05$); the ratings of the English vs. Korean groups were not statistically different ($p=.896$). There was also a significant Language effect for the [ɕ-ʃ] pair ($F(2,60)=5.810$, $p<.01$), though the z-scores from all three language groups were below 0, indicating that the two sounds were very similar to all the listeners. The significance came from the difference of the Mandarin vs. English groups ($p<.01$), and of the Korean vs. English groups ($p<.05$). There was no difference between the Mandarin vs. Korean groups ($p=.898$).

In the [a_a] context (Figure 2-6 (c)), the [s-ɕ] pair was rated as more different by the Mandarin listeners than by the other two groups, and a trend towards higher difference ratings from the English speakers than from the Korean speakers was present. Subsequent analyses showed that the factor Language, though not significant in the [s-ʃ] pair ($F(2,57)=2.483$, $p=.093$), was significant in the [s-ɕ] ($F(2,57)=18.642$, $p<.001$), and [ɕ-ʃ] pairs ($F(2,57)=5.187$, $p<.001$). Both of the significant effects came from the difference of the Mandarin vs. Korean groups, and of the Mandarin vs. English groups (all $p<.001$). No significant difference was found for the English vs. Korean groups ([s-ɕ]: $p=.752$; [ɕ-ʃ]: $p=.599$).

Note that in the [a_a] context, the Mandarin speakers rated the [ɕ-ʃ] pair as more different than the other two groups. A possible explanation for this is that Mandarin speakers usually perceive and adopt the rounding of the nonexistent alveo-palatal fricative [ʃ] (see section 1.3) to their native language as a front-rounded vowel [y] or glide [ɥ] (e.g., *Jo**sh*** →[tɕʰjaw**ɕy**]; Mi**ch**elle → [mi**ɕɥ**ɛɹ̩]). The more different ratings might be due to the fact that they perceived the rounding on [ʃ] as the front-rounded glide [ɥ]. In other words, the Mandarin listeners might have been comparing the similarity of [a**ɕ**a] and [a**ɕɥ**a], and as a consequence, they rated the fricative pair as more different than did the English and Korean listeners. The lower similarity ratings from the Korean and Mandarin groups on the [ɕ-ʃ] pair than the English group in the [i_i] context could be explained the same way. Korean speakers, like Mandarin speakers, usually perceive and adopt the rounding of the nonexistent [ʃ] to their native language as a labial glide [ʷ] (e.g., ***sh**e* →[**ɕʷ**i]; ***Sch**ick* → [**ɕʷ**ikʰɨ]) (Suh 2009a). In other words, the Korean listeners might have been comparing the similarity of [i**ɕ**i] and [i**ɕʷ**i], and thus rating the fricative pair as more different than did the English listeners. Korean listeners did not rate the [ɕ-ʃ] pair as more different than the English listeners in the [a_a] context presumably because in this vowel context, the nonexistent [ʃ] is perceived as [ɕ] (c.f., (c) and (d) in Table 1-4). Thus Korean listeners rated the [ɕ-ʃ] pair as very similar. This explanation also accounts for the fact that we only found the difference in non-rounding contexts, [i_i] and [a_a] (in both [i_i] and [a_a] contexts for the

Mandarin group, and in the [i_i] context for the Korean group), but not [u_u]. The statistical results are summarized in Table 2-11.

Table 2-11 Summary of similarity rating results by Vowel Context

[u_u]	***Simple effect of Language in [s-ʃ]	***Mandarin & Korean **English & Korean Mandarin & English (*p*=.459)
	***Simple effect of Language in [s-ɕ]	**Mandarin & Korean *English & Korean Mandarin & English (*p*=.542)
	Simple effect of Language in [ɕ-ʃ] (*p*=.542)	
[i_i]	Simple effect of Language in [s-ʃ] (*p*=.06)	
	*Simple effect of Language in [s-ɕ]	*Mandarin & English *Mandarin & Korean English & Korean (*p*=.896)
	**Simple effect of Language in [ɕ-ʃ]	**Mandarin & English *Korean & English Mandarin & Korean (*p*=.898)
[a_a]	Simple effect of Language in [s-ʃ] (*p*=.093)	
	***Simple effect of Language in [s-ɕ]	***Mandarin & Korean ***Mandarin & English English & Korean (*p*=.752)
	***Simple effect of Language in [ɕ-ʃ]	***Mandarin & Korean ***Mandarin & English English & Korean (*p*=.599)

2.3.3 General discussion

We expected that English listeners would judge *s* and *sh* (both the [s-ɕ] and [s-ʃ] pairs) as more different than Korean listeners, based on the different phonological relationships of the sound pairs in these two languages: in English, though participate in limited alternation, *s*/*sh* may signal differences in meaning and the choice of one vs. the other is not predictable from the environment, but in Korean, the occurrence of *s* vs. *sh* is predictable based on distribution, and the two sounds participate in regular and productive morphological alternations. The overall results suggest that, as noted in Johnson & Babel (2010), the phonological relationship affects the perceived phonetic similarity of two sounds. We found that the Korean listeners rated *s* and *sh* (both [s-ɕ] and [s-ʃ]) as more similar to each other than did the English listeners. This pattern was reliably present in the [u_u] context, and we observed a trend in this direction for the [a_a] and [i_i] contexts. The similar judgments of the [ɕ-ʃ] pair for the English group echo the findings in Lisker (2001) and McGuire (2007) that English listeners in general categorize the nonexistent sound in their native language, [ɕ], as /ʃ/.

Along the same lines, we should expect to see the English listener's judgments on the [s-ɕ] and [s-ʃ] pairs to be similar, since [ɕ] and [ʃ] are perceived as the same category (/ʃ/) (Lisker 2001; McGuire 2007). The prediction holds for the [a_a] and [u_u] contexts: the ratings of the [s-ɕ] pair were not significantly different from those of the [s-ʃ] pair in these contexts ([a_a]: $F(1,19)=1.514$, $p=.234$; [u_u]: $F(1,19)=1.197$, $p=.288$). However, the prediction does not hold for the [i_i] context ($F(1,19)=35.374$, $p<.001$): English listeners' judgments on the [s-ɕ] pair were significantly different from those on the [s-ʃ] pair. Kathleen Hall (personal communication) suggests that this pattern, whereby English listeners judged [s-ɕ] as more similar in the [i_i] context than in the other vowel contexts, might be explained as follows: English listeners, encountering the non-occurring sound [ɕ] in their native language, might have perceptually assimilated it to the category /ʃ/ in the [u_u] and [a_a] contexts (c.f., Lisker (2001) and McGuire (2007)), but treated [ɕ] as a positional variant of [s] in pre-palatal position, the [i_i] context (p.c. Kathleen Hall). This could be a possible explanation of why English and Korean groups appear to pattern together for the [s-ɕ] pair since the more similar ratings of this pair was driven by the [i_i] context.

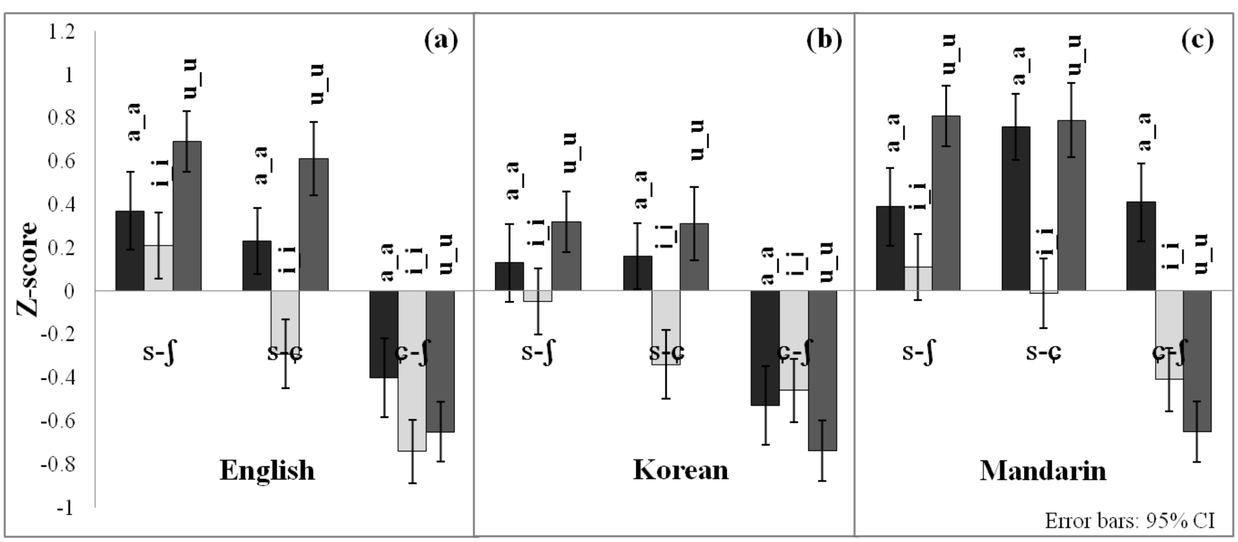
The results from the English and Korean groups have suggested that the similarity rating task, along with the previous studies employing the same task, did reflect the phonological relationships of sounds in their native language: English listeners judged *s/sh* as more different than did Korean listeners. We can now turn to the Mandarin results. If distributional predictability in the absence of morphological alternation is not sufficient to group two sounds as variants of the same phoneme, then we should expect to see the Mandarin listeners' ratings of *s/sh* to be more different than those of the Korean speakers (*Distribution Plus Alternation Hypothesis*). If distributional predictability alone is sufficient to group two sounds as variants of the same phoneme, then we should expect to see the Mandarin speakers' ratings to be comparable to those of the Korean speakers (*Distribution Alone Hypothesis*).

We found that the Mandarin listeners rated *s* and *sh* (both [s-ɕ] and [s-ʃ]) as more different from each other than did the Korean listeners, in all three vowel contexts (the palatalization context [i_i] as well as the other vowel contexts). Crucially, the Mandarin speakers rated the [s-ɕ] pair (the two fricatives that are in complementary distribution, as in Korean, but do not alternate) as significantly more different than did the speakers of Korean, in which the two fricatives *do* alternate. This suggests that Mandarin listeners, who are exposed to only distributional evidence, are less likely to group *s* and *sh* as variants of the same category than Korean listeners, who are exposed to both distributional evidence and morphological alternation.

One might suspect that the results could be explained not on the basis of phonological relationships but rather to different levels of English proficiency: perhaps the Mandarin speakers had established a *sh* category as a result of greater exposure to English. However, the Korean group had a slightly higher rating of their English ability (4.65/7.00) than the Mandarin group (4.6/7.00), and the Korean participants were recruited in the United States where English input is more abundant, whereas the Mandarin participants were recruited in Taiwan where English input is limited. If degree of exposure to English were a major factor driving the results, we should have seen the reverse bias: lower similarity ratings of *s* and *sh* in the Korean group than in the Mandarin group.

Of interest here are the by-vowel context results, as shown in Figure 2-7 ((a): English; (b): Korean; (c): Mandarin). The scores above zero means "more different" and the scores below zero means "more similar".

Figure 2-7 Similarity rating results by Vowel Context paneled by Language



The similarity judgments from the listeners of the same language group varied to different extents depending on the vowel contexts in different fricative pairs. Most importantly, though the Mandarin speakers' similarity judgments patterned overall with those of the English speakers—the English and Mandarin speakers rated *s* and *sh* as more different than did the Korean speakers—we found that the similarity judgments of the listeners on the *s-sh* pairs from all three languages varied according to the *vowel contexts*.

In the [i_i] context, the perceived perceptual similarity of the *s-sh* pairs ([s-ɕ] and [s-ʃ]) increased for the speakers of all three languages. It is not surprising that this context caused increased similarity for the Korean (Figure 2-7 (b)) and Mandarin listeners (Figure 2-7 (c)) since preceding the vowel *i* is the context where palatalization occurs in their native language. The increased similarity judgments from the Korean and Mandarin groups, though significantly more similar from the Korean group than from the Mandarin group (see section 2.3.2), suggest that both *distribution* and *alternation* are relevant. The increased similarity in the [i_i] context than in the other non-palatal vowel contexts ([u_u] and [a_a]) for the Mandarin group suggests an effect of *distribution* in the absence of alternations. That is, if alternation were the only factor in deciding sound membership, we should have seen a reduced phonetic distance between *s* and *sh* in the [i_i] context for the Korean group only, but not for the Mandarin group, since distribution would have been irrelevant. On the other hand, the degree of increased similarity being significantly less from the Mandarin group than from the Korean group in the [i_i] context, as well as in the other contexts, suggests an effect of *alternation*. That is, if distribution were the only factor in grouping sounds as members of the same category, we should have seen a similar degree of reduced phonetic distance between *s* and *sh* in the [i_i] context and in the other vowel contexts in the two language groups, since alternation would have been irrelevant.

Interestingly, the [i_i] context also caused increased perceptual similarity for the English listeners whose native language contrasts *s* and *sh* (Figure 2-7 (a)). In other words, we did not see the same degree of different ratings from the three vowel contexts even in a language in which the two sounds are contrastive in all these contexts and signal lexical differences (e.g., ***s**ee* vs. ***sh**e*, ***s**ue* vs. ***sh**oe*, ***s**ock* vs. ***sh**ock*). Instead, we found more similar judgments by the English listeners in the [i_i] context than in the [u_u] and [a_a] contexts. One possible explanation for this is that in the pre-palatal context (before the high-front vowel [i]), the place of articulation of the dental *s* is made more palatal, and thus the phonetic distance between *s* and the palatal *sh* is reduced. In other words, even in a language such as English, in which *s* and *sh* are contrastive in pre-palatal context, the phonetic distance between *s* and *sh* is still reduced. Another explanation for the increased perceptual similarity in English is that *s*, in connected speech, alternates with *sh* in pre-palatal contexts (e.g., *miss* [mɪs] ~ *miss you* [mɪʃju]; c.f., discussion in section 1.3.1). The phonetic distance between *s* and *sh* might be reduced because of the alternation in the [i_i] context. In other words, the reduced perceptual distance could be due to the effect of phonetic alternation.

One might wonder whether the increased perceptual similarity in the [i_i] context from the English group was due to the morphological alternations discussed in section 1.3.1. This possibility is unlikely since the morphological alternations between *s* and *sh* are limited to certain suffixes that do not necessarily provide the pre-palatal context. In other words, these alternations are morphologically conditioned, and not necessarily depending on phonological environment. If the increased perceptual distance were due to the morphological alternations, we should have observed a similar effect in other vowel contexts.

In the case of Mandarin, the lack of morphological alternation of the two sounds seems to be taking precedence over the predictability of the phonological environments (indicated by the overall more different ratings of *s/sh* from the Mandarin group than from the Korean group). This is interesting given that alternation does not seem to account for the English results in the same experiment (indicated by the overall more different ratings from the English group than from the Korean group). The limited morphological alternation discussed above might be the reason why the unpredictability of the two sounds from the phonological environments seems to be taking precedence in the English case. For discussion of possible future research that might tease apart the weighting of different factors in sound memberships, see section 5.4.

To summarize the results so far, we found overall higher difference ratings on the similarity judgment tasks from the English and Mandarin listeners than from the Korean listeners. This finding suggests that speakers who have access only to distributional evidence (*s/sh* in Mandarin) are less likely to analyze sounds as members of a single category than speakers who are exposed to evidence from both distribution and morphological alternation (*s/sh* in Korean), supporting the *Distribution Plus Alternation Hypothesis*. We also found that the similarity ratings of the pre-palatal context ([i_i]) were significantly lower for the Korean group than for the Mandarin group, suggesting that distribution reinforced by alternation produced a stronger motivation for learners to group sounds to the same category than distribution alone.

Furthermore, we found that similarity judgments varied depending on the vowel contexts for listeners of all three languages. The varying similarity judgments according to the vowel contexts in all three language groups cannot be explained by a *categorical* view of phonological relationships. From the point of view of categorical phonological relationships, two sounds are either variants of the same category or surface forms of separate categories. For example, we would expect English speakers to judge the similarity of *s* and *sh* the same in all vowel contexts in which the two sounds may contrast. The same should be true for Korean, where the choice of *s* vs. *sh* is predictable from the environment. The fact that we saw varying judgments according to vowel context (c.f., Figure 2-6 and Figure 2-7) suggests that the judgment of two sounds as members of the same or separate categories is not an all-or-nothing judgment/mapping.

2.4 Summary

The set of experiments in this chapter investigated how listeners of English, Korean and Mandarin rated the similarity of *s* and *sh*, two sounds that participate in different phonological relationships in these languages. The results from the English and Korean groups showed that the different relationships were reflected in their similarity judgments. As expected, the Korean listeners, in whose language *s* and *sh* are in complementary distribution and participate in productive morphological alternations, rated these sounds as more similar than did the English listeners, in whose language *s* and *sh* are not predictable from the phonological environment. The similarity judgments of the Mandarin group, in which *s* and *sh* show distributional predictability but do not participate in morphological alternations, resembled those of the English group rather than the Korean group. Consequently, the results support the hypothesis that *alternation* reinforces the mapping of two sounds to the same category, giving a stronger effect than distribution alone is (*Distribution Plus Alternation Hypothesis*). However, we also found that the similarity judgments from the listeners of all three languages *varied* depending on the vowel contexts. The results suggest that both distribution and alternation contribute to the determination of sound category memberships, and that the judgment of two sounds as members of the same or separate categories is not necessarily absolute. The next chapter presents the results of another probe, discrimination on a continuum, to investigate the behavior of English, Korean, and Mandarin speakers with respect to the *s/sh* sound difference.

Chapter 3 Discrimination on a Continuum

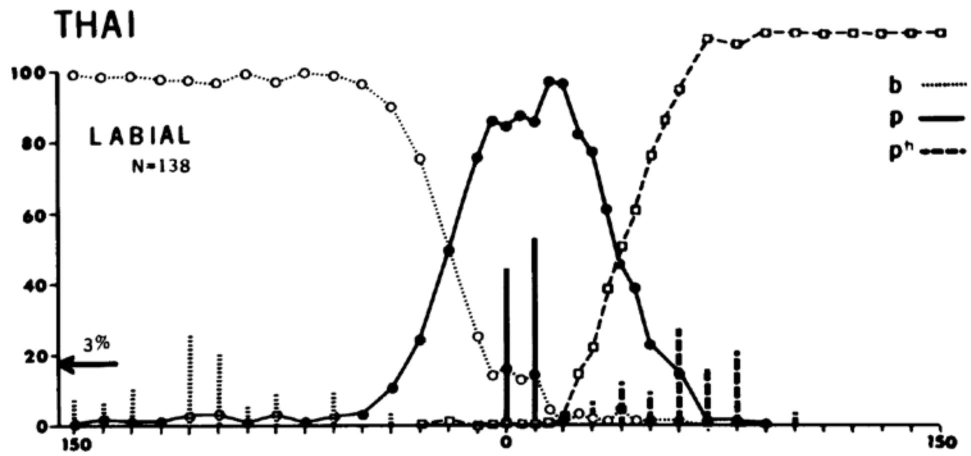
This chapter presents the results of a discrimination experiment designed to investigate how English, Korean, and Mandarin speakers perceive pairs of sounds on the *s-sh* continuum. The prediction of the distributionally based definition of phonological relationships is that phonetically similar sounds in complementary distribution are variants of the same category, and that Mandarin *s/sh*, like Korean *s/sh*, should be analyzed as variants of the same category (*Distribution Alone Hypothesis*). However, if distribution alone is not assumed to force learners to map sounds in complementary distribution onto the same underlying category, Mandarin *s/sh* should be analyzed as separate categories (*Distribution Plus Alternation Hypothesis*). The results from the first study (Chapter 2), in which English and Mandarin speakers patterned together in contrast to Korean speakers, gave support to the *Distribution Plus Alternation Hypothesis*. In this study, I will show that the accuracy results did not yield a difference in the ability to discriminate the two sounds according to language background, contradicting the results of the first study. Furthermore, the response time results showed that the English speakers patterned with the Korean speakers in that, overall they took less time than did the Mandarin speakers in their discrimination of *s/sh*. Possible explanations for the discrepancy between the results of the similarity rating and discrimination experiments will be discussed.

The structure of this chapter is as follows. Section 3.1 gives a brief review of the previous literature of discrimination on a continuum. Section 3.2 presents the methodology of the experiment and section 3.3 gives the results, followed by a discussion in section 3.4. Section 3.5 is a summary.

3.1 Introduction

Studies of identification and discrimination of sounds varying by equal intervals on a continuum have revealed that speakers can more easily discriminate and identify two sounds that fall across a phoneme boundary of their native language than two sounds within a phoneme category (categorical perception; e.g., Lisker & Abramson 1970; Lasky et al. 1975; MacKain et al. 1981; Best et al. 1988; Werker & Lalonde 1988; Kuhl 1991; Lisker 2001; Kazanina et al. 2006). Lisker & Abramson (1970) asked Thai, English, and Spanish speakers to *identify* stimuli from continua manipulating voice onset time (VOT) for three different places of articulation (labial [b]-[p^h], apical [d]-[t^h], and velar [g]-[k^h]), using their native language orthography. Thai has a three-way laryngeal contrast in stops (voiced, unaspirated voiceless, aspirated; /**b**aa/ 'crazy', /**p**aa/ 'aunt', vs. /**p**$^{\mathbf{h}}$aa/ 'cloth') while both English and Spanish have a two-way laryngeal contrast (English: ***b****at* vs. ***p****at*; Spanish: ***p****ata* 'leg' vs. ***b****ata* 'bath robe'). Figure 3-1, Figure 3-2, and Figure 3-3 provide the *labial* voicing continuum results from Lisker & Abramson's Thai, English, and Spanish listeners, respectively, as examples. The *x* axis represents voice onset time in milliseconds and the *y* axis represents the percent of responses identifying a particular stimulus as member of the relevant category. The bars show frequency distributions of VOT values measured in real speech.

Figure 3-1 Thai listeners' identification on [ba]-[pʰa] continuum (Lisker & Abramson 1970: 14): 0 = the release of the constriction



On the [b]-[pʰ] continuum, in which VOT ranged from -150 ms to 150 ms, the results from the Thai listeners showed *two* places where responses changed abruptly, -20 ms VOT and 40 ms VOT. The /b/ responses dropped down around -20 ms VOT, and /p/ responses increased. After 40 ms VOT, /p/ responses dropped down and /pʰ/ responses started to increase. In other words, the identification results on the VOT continuum from the Thai listeners showed evidence of two category boundaries, consistent with the Thai three-way contrast of stops (voiced, unaspirated voiceless, aspirated).

When English speakers and Spanish speakers were presented with the same continuum, however, the results showed evidence for only two categories, as shown in Figure 3-2 and Figure 3-3 respectively, consistent with the English and Spanish two-way contrast of stops. Spanish speakers identified sounds with negative VOT values as /b/, and /p/ responses increased after 0 ms VOT, while the cut-off point for English speakers was around 10-20 ms.

Figure 3-2 English listeners' identification on [ba]-[pʰa] continuum (Lisker & Abramson 1970: 13)

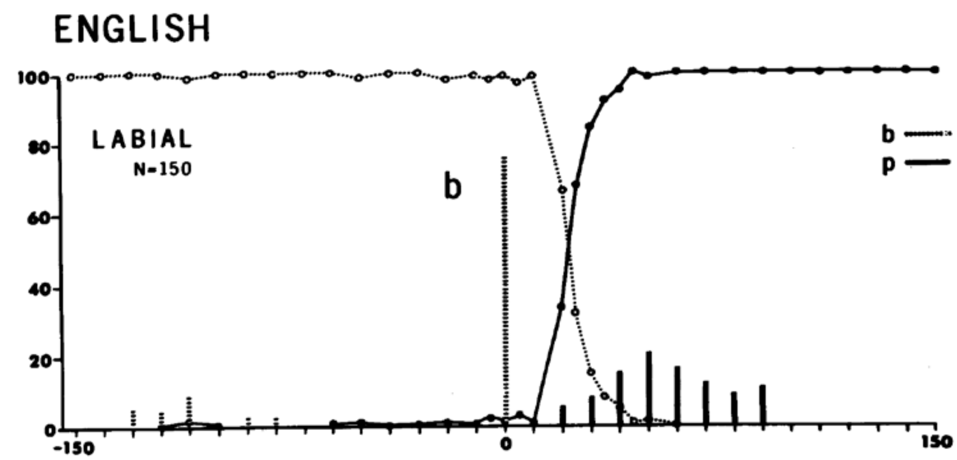
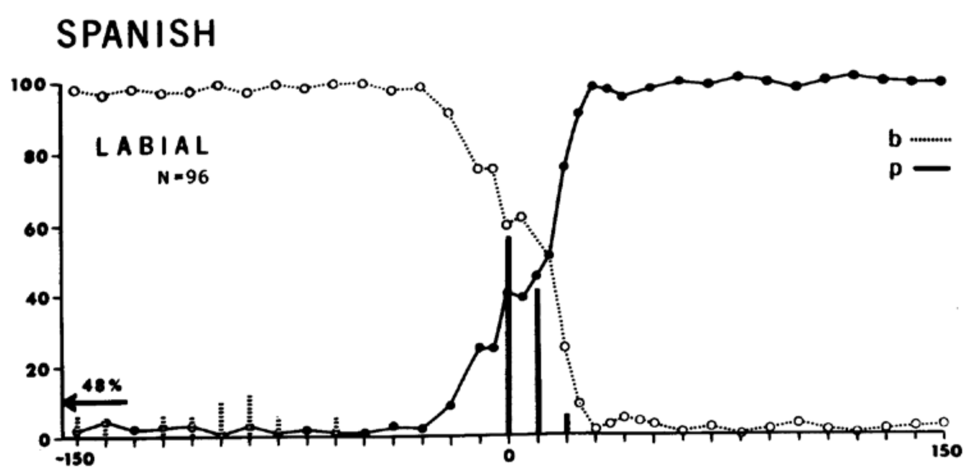
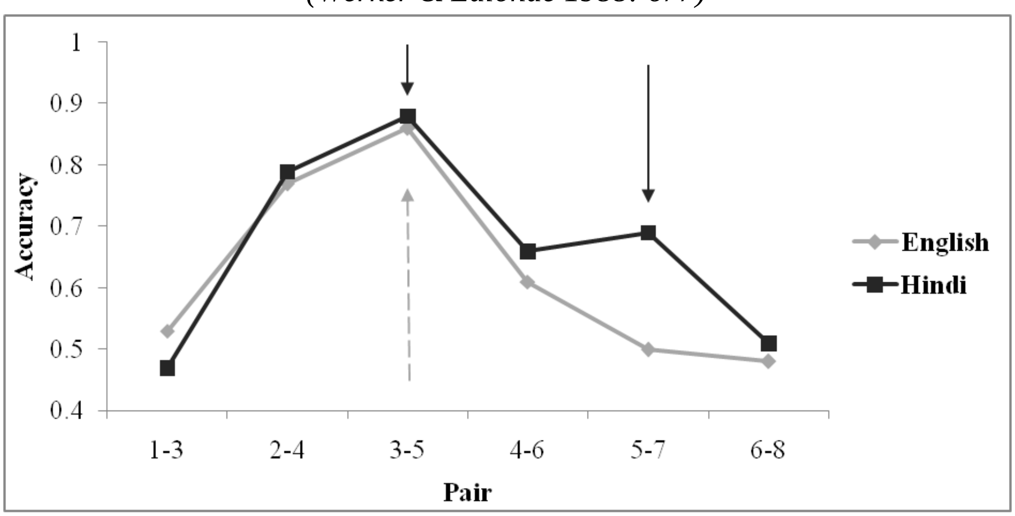


Figure 3-3 Spanish listeners' identification on [ba]-[pʰa] continuum (Lisker & Abramson 1970: 12)



Categorical perception corresponding to phoneme boundaries is also reflected in the accuracy of *discrimination* by different language groups. For example, Werker & Lalonde (1988) synthesized an 8-step continuum from [ba] to [ɖa] manipulating place of articulation (voiced labial stop to voiced retroflex), and played sound pairs two steps apart on the continuum in an ABX paradigm to Hindi speakers and English speakers. Hindi contrasts stops at three places of articulation—labial, alveolar, and retroflex—while English contrasts only two within this range—labial and alveolar. Hindi speakers' discrimination on the continuum from [ba] to [ɖa] (represented by square markers), as in Figure 3-4 (redrawn from Table 1 in Werker & Lalonde 1988), showed two peaks of accuracy. The two peaks, indicated by the solid arrows, matched the Hindi three-way contrast in place of articulation (labial, alveolar, retroflex). English speakers' discrimination on the same continuum (represented by diamond markers), however, only showed one accuracy peak, indicated by the dotted arrow, matching the English two-way contrast in place of articulation (labial, alveolar) on the continuum (Werker & Lalonde 1988).

Figure 3-4 Discrimination on an eight-step continuum from [ba] to [ɖa] (Werker & Lalonde 1988: 677)



Along the same lines, in Beckman & Pierrehumbert's (2000) perceptual experiment in *identifying* naturally produced Korean syllables [ɕi], [ɕ'i], [sɨ] and [s'ɨ] described in section 1.2, English listeners showed categorical perception of *s* and *sh*, corresponding to the phonemic status of *s/sh* in English, while Korean listeners showed only chance level perception in identifying the two sounds, corresponding to the non-contrastive status of *s*/*sh* in their native language.

The goal of the current experiment was to determine whether speakers whose native language offers different types of evidence for the relationship between *s* and *sh* exhibit differences in the categorization of sounds along a *s*-*sh* continuum.

3.2 Methodology

An eight-step synthesized continuum from *s* to *sh* was presented to native speakers of Mandarin, English, and Korean. Another eight-step continuum from *f* to *s* was synthesized as a control. The two sounds *f* and *s* signal lexical differences in English and Mandarin (e.g., minimal pair ***f**ee* vs. ***s**ee* in English, and [sǎ] 'spill' vs. [fǎ] 'hair' in Mandarin). Though *f* does not exist in the Korean consonant inventory (c.f., Table 1-3), it is often categorized as /p^h/ (Kim 2009: 170). Therefore, if the Korean speakers categorized [f] as [p^h] on the control *f*-*s* continuum, we expect a categorical judgment from speakers of all three languages, since /p^h/ and /s/ contrast in Korean (e.g., /p^hal/ 'arm' vs. /sal/ 'flesh', and /p^hul/ 'grass' vs. /sul/ 'liquor') and /f/ and /s/ contrast in the other two languages.

In this experiment, instead of using full syllables as stimuli, as in most previous discrimination studies, only the frication portion was used. The reason for this decision was that since *s* and *sh* occur in distinct phonological environments in Korean and Mandarin, it is impossible to put the fricatives in the same vowel context. Furthermore, I tested only *discrimination*, not identification. The reason for this was that, in Korean, these two sounds are represented with the same orthographic symbol 'ㅅ'. Thus, there is no way to represent the *s-sh* distinction in Korean. I will discuss the possible role of orthography in Chapter 5.

3.2.1 Participants

20 participants from each language group were recruited for this set of experiments. Participants in the English group (11 male, 9 female, aged 18-22), all monolingual speakers of English, were recruited at Stony Brook University, and received course credit for their participation. Although 22 English participants were originally recruited for this experiment, two participants were excluded due to their high no-response rate (1 female, 41% and 1 male, 46%; the no-response rate of the included participants was 1.3% in the English group). Participants in the Korean groups (6 male, 14 female, aged 18-38), all native speakers of Korean from South Korea, were recruited at Stony Brook University, and received payment for their participation. To estimate possible influence from English, participants were asked to rate their English ability; the average rating was 4.65 on a 7-point scale. They all received up to a high school education in South Korea before coming to Stony Brook University for undergraduate or graduate education. Participants in the Mandarin group (1 male, 19 female, aged 20-22) were all native speakers of Taiwanese Mandarin, and were recruited at National Chiao Tung University in Taiwan for course credit or payment. Their average self-rating of English ability was 4.4 on a 7-point scale. None reported any hearing deficiencies.

3.2.2 Design and materials

In order to present the same stimuli to all three language groups, the same endpoint stimuli, [s] and [ɕ], were used. Although English [ʃ] differs acoustically from Mandarin [ɕ] (Ladefoged & Maddieson 1996; Li 2008), English speakers generally categorize [ɕ] as /ʃ/ (c.f. the results from similarity ratings in Chapter 2; Lisker 2001; McGuire 2007). The endpoints [s] and [ɕ] were spliced from [si] and [ɕi] syllables spoken by a trained female phonetician whose native language is Mandarin, using Praat software package (Boersma 2001). The Mandarin speaker was chosen to record the stimuli because she was able to produce the syllables [ɕi] natively and [si] from extensive English exposure and professional training.

The endpoints were synthesized proportionally to create an eight-step continuum using Audacity (http://audacity.sourceforge.net/), following the methodology of Suh (2009a).[10] Step 1 was 100% [s], step 2 was 85.7% [s] and 14.3% [ɕ], and additional steps were synthesized as in Table 3-1.

[10] The continuum was not produced by manipulating a certain acoustic dimension of fricatives. Rather, the continuum was produced by using different proportions of the end points [s] and [ɕ] by overlapping different numbers of [s] sound track and [ɕ] sound track. Step 1 was created by overlap 7 tracks of the endpoint [s], and step 2 was created by overlapping 1 track of endpoint [ɕ] with 6 tracks of endpoint [s], etc.

Table 3-1 Eight-step continuum from [s] to [ɕ]

Step	**Stimuli**
1	100% of [s]
2	85.7% of [s] and 14.3% of [ɕ]
3	71.4% of [s] and 28.6% of [ɕ]
4	57% of [s] and 43% of [ɕ]
5	42.7% of [s] and 57.3% [ɕ]
6	28.6% of [s] and 71.4% [ɕ]
7	14.3% [s] and 85.7% of [ɕ]
8	100% of [ɕ]

The length of each stimulus was 270ms, a rough average of all the endpoint stimuli. The intensity of the stimuli was scaled to 56 dB, the averaged intensity of the endpoints, using Praat software. The *f-s* continuum was synthesized in the same way.

Six pairs consisting of sounds two steps apart were presented in an ABX design (6 pairs x 4 orders, ABB, ABA, BAA, BAB = 24) for each continuum (24 x 2 continua = 48) with an ISI of 500 ms. Order was randomized for each participant using E-Prime software (v2.0; Psychological Software Tools, Pittsburgh, PA). Listeners heard each of the ABX trials (48 trials) twice in each of the 2 blocks (48 x 2 repetitions x 2 blocks = 192).

The experiment was a two-factorial design, as in Table 3-2. Participants compared and discriminated six pairs of sounds two steps apart.

Table 3-2 Discrimination task design

Between-subject factor	Language	Mandarin, English, Korean
Within-subject factor	Pairs (every two-step apart pairing)	*s*-----------------------------------*sh* steps 1-3, 2-4, 3-5, 4-6, 5-7, 6-8 *f*------------------------------------*s* steps 1-3, 2-4, 3-5, 4-6, 5-7, 6-8
Dependent variable	Accuracy Response times	

3.2.3 Procedures

The participants took part in the experiments individually or in groups of up to four in separate booths, using a computer connected to a keyboard with two keys labeled '1' and '2'.[11] Participants were presented with written instructions on the computer screen in their native language saying that they would hear three sounds per trial, and they should indicate whether the third sound was the same as the first sound or the second sound by pressing the keys ('1' or '2') on the keyboard. No feedback was given. There were two blocks for the experiment with a break between the blocks. Participants had 4000 ms to respond before the next trial started if they did not respond to a given trial. All stimuli were presented binaurally over headphones at a comfortable listening level. The participants completed a 10-trial practice randomly chosen from the test trials, and had the opportunity to ask questions before proceeding to the experiment. The experiment lasted approximately 10 minutes.

[11] The labels '1' and '2' were put on the keys 'd' and 'l' on a keyboard because of their relative central position on the keyboard.

3.3 Results

Listeners from all three language groups were expected to show evidence of a category boundary for the *f*-*s* continuum, because *f* and *s* contrast in English and Mandarin. Although *f* does not exist in Korean, it is perceived and categorized as $/p^h/$ by Korean speakers (Kim 2009: 170). If listeners show evidence of a category boundary only for the *f*-*s* continuum, and not for the *s*-*sh* continuum, it suggests that *s* and *sh* are perceived as variants of the same category. If their results show evidence of a boundary for both continua, this suggests that *s* and *sh*, just like *s* and *f*, are perceived as separate categories. In other words, if listeners show similar discrimination patterns for the *s*-*sh* and the *f*-*s* continua, then *s* and *sh* should be considered different categories. Most importantly, if a category boundary is present in the Mandarin as well as in the English listeners' discrimination on the *s*-*sh* continuum, but not in the discrimination of the Korean listeners, this suggests that *alternation* plays a crucial role in defining phonological relationships, because although the distributional patterns of these two sounds are similar in Mandarin and Korean, the two sounds alternate in Korean but do not in Mandarin. The predictions are summarized in Table 3-3 (√ = categorical boundary).

Table 3-3 Predictions of discrimination on a continuum

Distribution Alone Hypothesis: For Mandarin and Korean speakers, *s* and *sh* are single category.

Predictions for discrimination results:

	f-s **continuum**	***s-sh*** **continuum**
English	√	√
Korean	√	
Mandarin	√	

Distribution Plus Alternation Hypothesis: For Korean speakers, but *not* Mandarin speakers, *s* and *sh* are single category.

Predictions for discrimination results:

	f-s **continuum**	***s-sh*** **continuum**
English	√	√
Korean	√	
Mandarin	√	√

The accuracy results for the three language groups on the two continua are shown in Table 3-4 and Table 3-5, with standard deviations in parentheses, and represented in graphs in Figure 3-5 and Figure 3-6, respectively.

Table 3-4 Discrimination accuracy results on f-s continuum

Language \ Pairs	**1-3**	**2-4**	**3-5**	**4-6**	**5-7**	**6-8**
Mandarin	0.70 (0.14)	0.83 (0.12)	0.82 (0.14)	0.68 (0.18)	0.56 (0.14)	0.47 (0.13)
English	0.77 (0.14)	0.79 (0.16)	0.81 (0.12)	0.69 (0.14)	0.51 (0.11)	0.44 (0.11)
Korean	0.65 (0.12)	0.85 (0.14)	0.81 (0.17)	0.64 (0.14)	0.50 (0.11)	0.48 (0.07)

Table 3-5 Discrimination accuracy results on s-sh continuum

Language \ Pairs	**1-3**	**2-4**	**3-5**	**4-6**	**5-7**	**6-8**
Mandarin	0.55 (0.15)	0.75 (0.16)	0.78 (0.17)	0.68 (0.16)	0.58 (0.13)	0.55 (0.14)
English	0.57 (0.11)	0.76 (0.11)	0.84 (0.09)	0.71 (0.13)	0.60 (0.1)	0.51 (0.12)
Korean	0.59 (0.09)	0.73 (0.11)	0.83 (0.14)	0.75 (0.11)	0.62 (0.12)	0.49 (0.13)

Figure 3-5 Discrimination accuracy results on f-s continuum

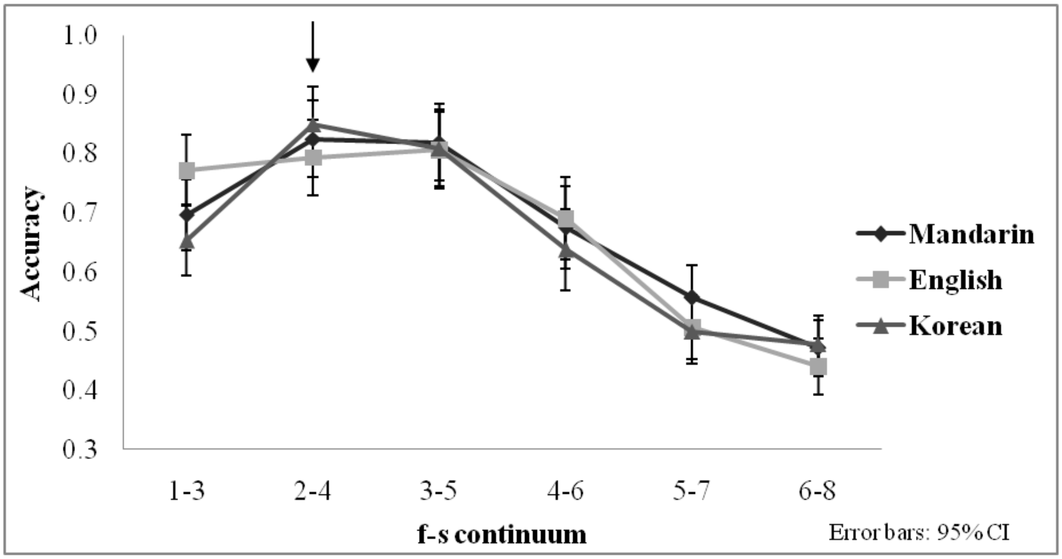
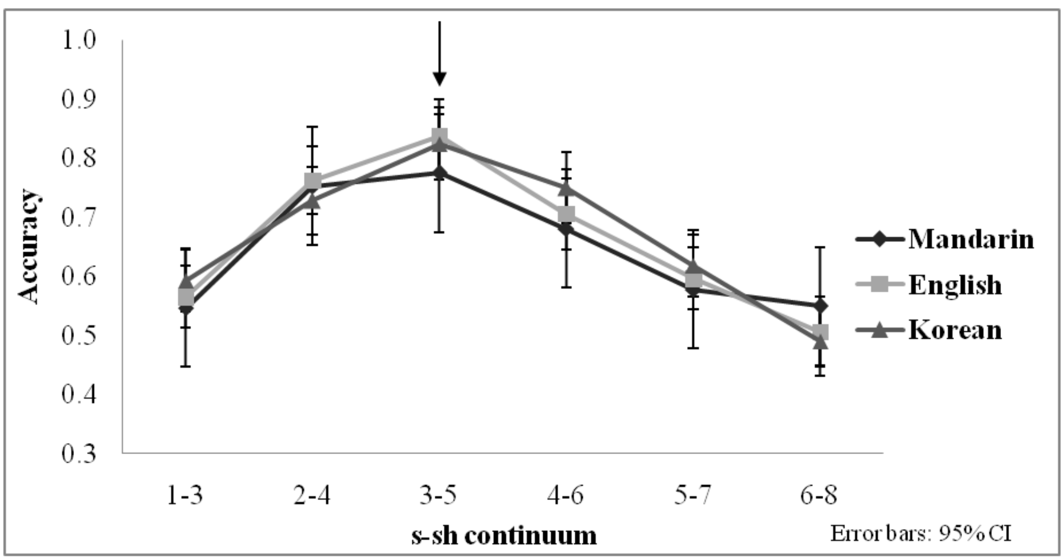


Figure 3-6 Discrimination accuracy results on s-sh continuum



In the figures, the *x* axis represents fricative pairs two steps apart on the continuum, and the *y* axis represents the percent of accurate responses. The Mandarin group is represented with diamond markers, English with square markers, and Korean with triangle markers.[12] As we can see from Figure 3-5 and Figure 3-6, all three groups showed evidence of a boundary on *both* continua (indicated by the arrows), located somewhere between steps 2-5 on the *f-s* continuum, and between steps 2-6 on the *s-sh* continuum.

A repeated-measures ANOVA (**Language** [Mandarin, English, Korean] **x Pair** [1-3, 2-4, 3-5, 4-6, 5-7, 6-8]) was performed to interpret the results for the *f-s* continuum. There was a main effect of Pair ($F(5,285)=87.381$, $p<.001$), but not of Language ($F(2,57)=.344$, $p=.710$). The results are indicative of the difference in accuracy across sound pairs, but not across the language groups. The interaction of Language and Pair was also not significant ($F(10,285)=1.583$, $p=.111$). In other words, the accuracy for a given pair was not statistically different depending on the native language of the participants.

Repeated-measures ANOVA on the *s-sh* continuum showed very similar results. There was a main effect of Pair ($F(5,285)=53.689$, $p<.001$), but not of Language ($F(2,57)=.553$, $p=.578$). The results again are indicative of a difference in accuracy across the sound pairs, but not across the language groups. The interaction of Language and Pair was not significant ($F(10,285)=1.037$, $p=.412$). In other words, the accuracy of a given pair was not statistically

[12] The assumption of normality was met with accuracy results. See Appendix E.2.

different depending on the language group of the participants. A summary of the statistical results is given in Table 3-6 (*: $p<.05$; **: $p<.01$; ***: $p<.001$).

Table 3-6 Summary of discrimination accuracy results

***f-s* continuum**	***s-sh* continuum**
Language	Language
***Pair	***Pair
Language x Pair	Language x Pair

To summarize, the accuracy results did not show a difference in the ability to discriminate *s* and *sh* according to language background. Instead, all three language groups showed better discriminability at one point on the *s-sh* continuum, parallel to discriminability along the *f-s* continuum on which the two sounds are considered to be separate categories in their native language.

3.4 Discussion

Going back to the predictions, repeated here in Table 3-7, we predicted that if listeners analyzed *s* and *sh* (like *f* and *s*) as different categories, then they should show similar discrimination patterns for the *s-sh* and the *f-s* continua. Most importantly, if a category boundary is present in the English listeners' discrimination as well as in the Mandarin listener's discrimination on the *s-sh* continuum, but not in the discrimination of the Korean listeners, this suggests that distributional criterion alone is not sufficient to lead learners to group sounds as variants of the same category.

Table 3-7 Predictions and actual results of discrimination on a continuum (√=categorical boundary)

Distribution Alone Hypothesis:

	***f-s* continuum**	***s-sh* continuum**
English	√	√
Korean	√	
Mandarin	√	

Distribution Plus Alternation Hypothesis:

	***f-s* continuum**	***s-sh* continuum**
English	√	√
Korean	√	
Mandarin	√	√

Actual results:

	***f-s* continuum**	***s-sh* continuum**
English	√	√
Korean	√	√
Mandarin	√	√

The perception on the *f-s* continuum for all the three languages was as expected, as these two sounds are distinctive for all the listeners. The results of the *s-sh* continuum, however, did not reveal a language effect. This finding is surprising given the findings of previous research (Beckman & Pierrehumbert 2000) in which English participants, in an identification task, showed categorical perception of *s* and *sh* while Korean participants showed only chance level perception, corresponding to the non-phonemic status of *s* and *sh* in their native language. These findings would have led us to expect a language effect on discrimination of the *s-sh* continuum by the English and Korean groups. We did not expect evidence of a discrimination discontinuity for the Korean listeners since *s* and *sh* are variants of the same category in their native language, as suggested in Beckman & Pierrehumbert (2000). Furthermore, the results seem to contradict the results of the first study (Chapter 2), in which English and Mandarin speakers patterned together in contrast to Korean speakers. All three language groups, on the other hand, showed better perception between steps 2-5, and chance-level perception closer to the end-points, suggesting a category boundary for all three language groups.

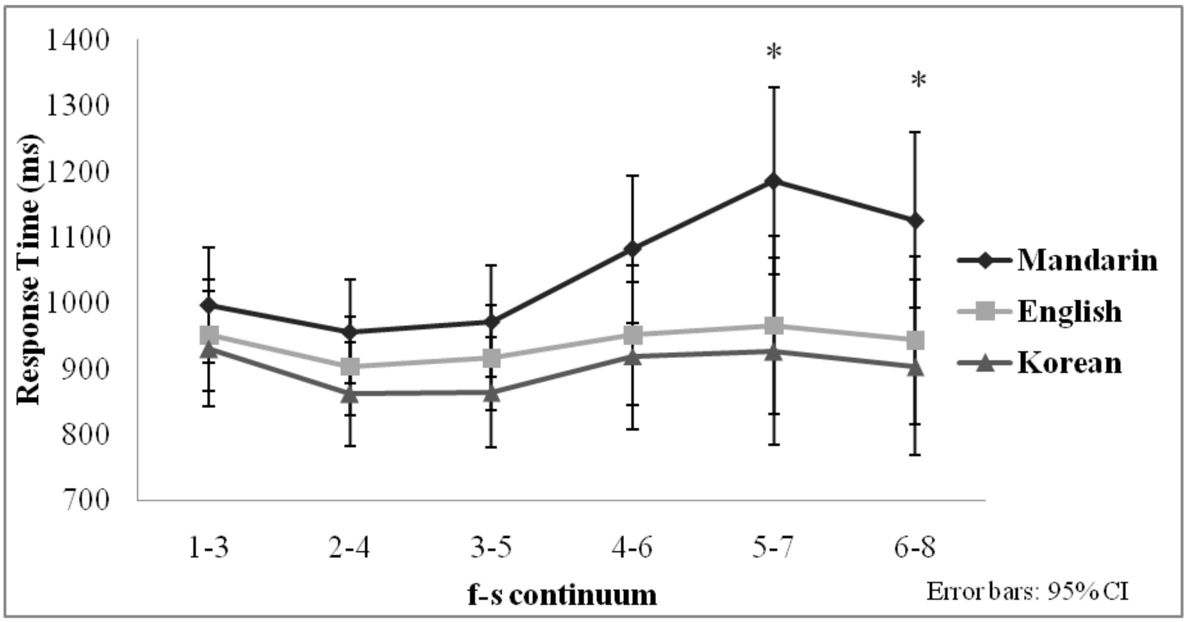
Several factors might explain the contradiction between these results and earlier research. In this experiment, only the frication portion (with possible leftover vowel transition) from the original C*i* context was used in synthesizing the continuum (c.f. section 3.2.2). Thus, it seems likely that this task tapped into *low-level auditory/acoustic* perception. In the previous studies, the feature being manipulated on the continuum (e.g., VOT in Lisker & Abramson (1970) and Pisoni & Tash (1974), place of articulation in Ganong (1980)) was put in a context that encouraged *phonological* perception. For example, in Lisker & Abramson (1970), a continuum of stops manipulating VOT in the context of CV (e.g., [ba] to [p^ha]) was played to English listeners and Thai listeners. In the context of (stressed) syllable onset position, English listeners rarely encounter unaspirated voiceless stops. In other words, English listeners only encounter two categories of stops in the given linguistic environment. Unaspirated voiceless stops appear allophonically in other contexts (e.g., after /s/, as in ***s*p*eak*** or in unstressed syllable onsets, as in *ha***pp***y*). English speakers might have discriminated all three allophonic variants if the context had been one that does not encourage phonological perception. Previous studies have shown that perception is influenced by native language phonotactic constraints (Massaro & Cohen 1983). In the case of VOT in the context of CV (Lisker & Abramson 1970), English listeners may have been biased/guided by their native phonotactic knowledge in terms of the aspiration environments. Another example of perception biased/guided by native language knowledge is repeated in Ganong (1980). Ganong uses a continuum of stops manipulating VOT in the context of a word or a non-word (e.g., *task*-**dask*, **tash*-*dash*), and the boundary of the categorical perception was biased towards the endpoint that was a word. In other words, on a *task*-**dask* continuum, listeners had more *task* responses while on the **tash*-*dash* continuum, listeners had more *dash* responses. In this linguistic context, listeners were biased/guided by their knowledge of the lexicon (Beckman & Pierrehumbert 2004; Hay et al. 2004; see also the discussion in section 4.5).

The experiment conducted here, on the other hand, included *no biasing context* and no lexical information. Thus, it seems likely that this task tapped into *low-level auditory/acoustic* perception rather than *phonological* processing. This explanation gained support from the response time (RT) results.

Previous research shows that short response latencies (as short as 500ms from acoustic onsets) are not significantly affected by linguistic knowledge (Fox 1984; Johnson & Babel 2010), and that response time increases as a positive function of uncertainty (Pisoni & Tash 1974): when the two sounds were across a category boundary, the RTs were shorter; when the two sounds fell within a category, two sounds that presumably pose more difficulty for the listeners, the RTs were longer. In other words, shorter responses to auditory stimuli were hypothesized to by-pass a higher level of conscious introspection in which RTs are argued to reflect uncertainty. Although this experiment was not designed in a way that encouraged speeded response, a post-hoc observation of the RT results reveals interesting patterns among the three language groups.

The RT results for the *f-s* continuum are shown in Figure 3-7.[13] The *x* axis represents the fricative pairs two steps apart on the continuum and the *y* axis represents the response times in milliseconds. The Mandarin group is represented with diamond markers, English with square markers, and Korean with triangle markers (*: p<.05).

Figure 3-7 f-s continuum response time

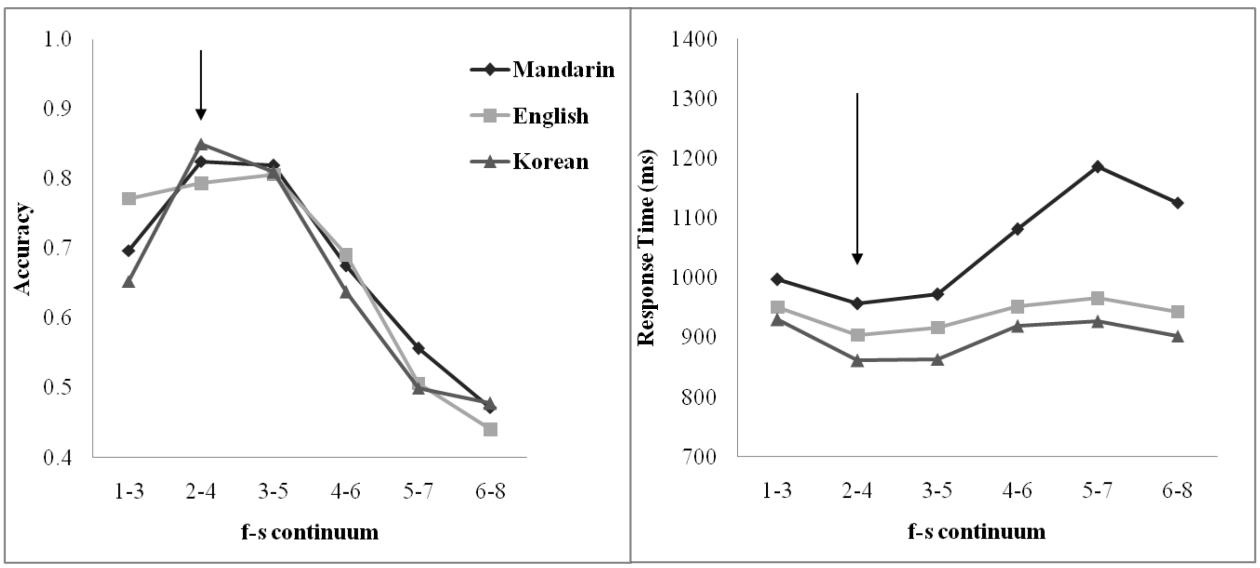


Overall, we see shorter response times from the English (mean: 947 ms) and Korean groups (mean: 900 ms) than from the Mandarin group (mean: 1053 ms), and the Mandarin speakers took longer to respond to stimuli towards the end of the continuum (though the same trend is present for both English and Korean groups).

[13] The assumption of normality was met with the RT results. See Appendix E.3.

A follow-up repeated-measures ANOVA showed a main effect of Pair ($F(5,285)=6.698$, $p<.001$) and a significant interaction of Language and Pair ($F(10,285)=2.043$, $p<.05$), indicating that the RTs were different among pairs depending on the language groups. The interaction was driven by the longer response times from the Mandarin group from steps 5-8 (Language effect in steps 5-7 and 6-8, both $p<.05$) towards the end of the continuum (all the other pairs, $p>.1$). An interesting pattern was revealed when we put the accuracy and the response time results side by side, as in Figure 3-8.

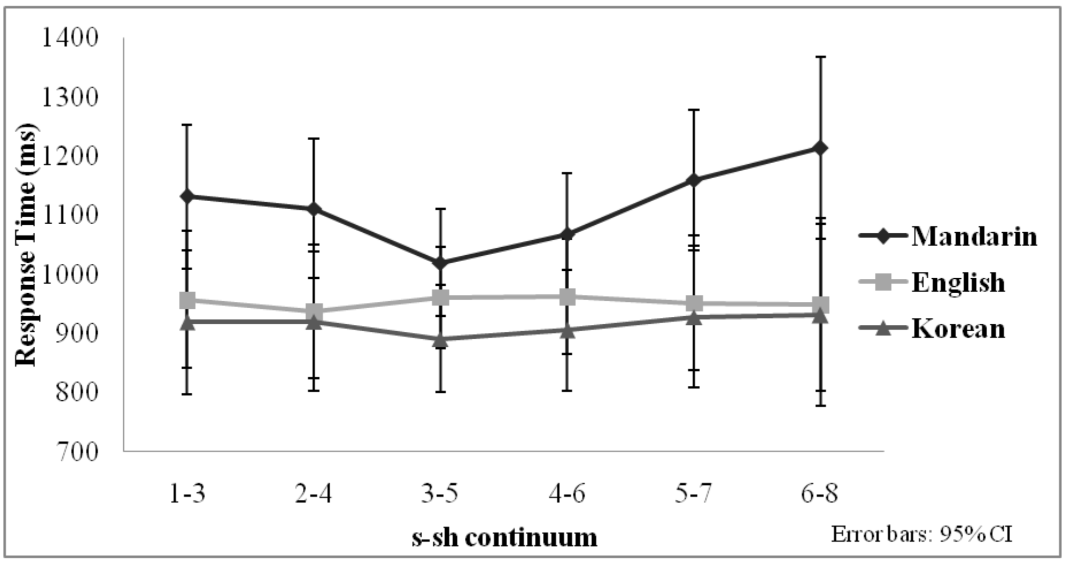
Figure 3-8 f-s continuum accuracy and response time results



The valley of the RT results in the right panel, indicated by an arrow, corresponded nicely to the peak of the accuracy on the left panel. This trend on the RT results seems to be present in all three language groups, but is only significant in the Mandarin group (linear trend, $p<.01$). In other words, only the Mandarin speakers took significantly less time responding to the pairs that they perceived more accurately, while they took longer to respond to the pairs they processed with lower accuracy.

Along the same lines, we observe similar RT patterns on the *s-sh* continuum, as shown in Figure 3-9. Both the English and Korean speakers took less time in responding (English: 967 ms; Korean: 916 ms; Mandarin: 1117 ms), and the RTs were not different depending on pairs. On the other hand, the RTs for the Mandarin group showed a quadratic trend.

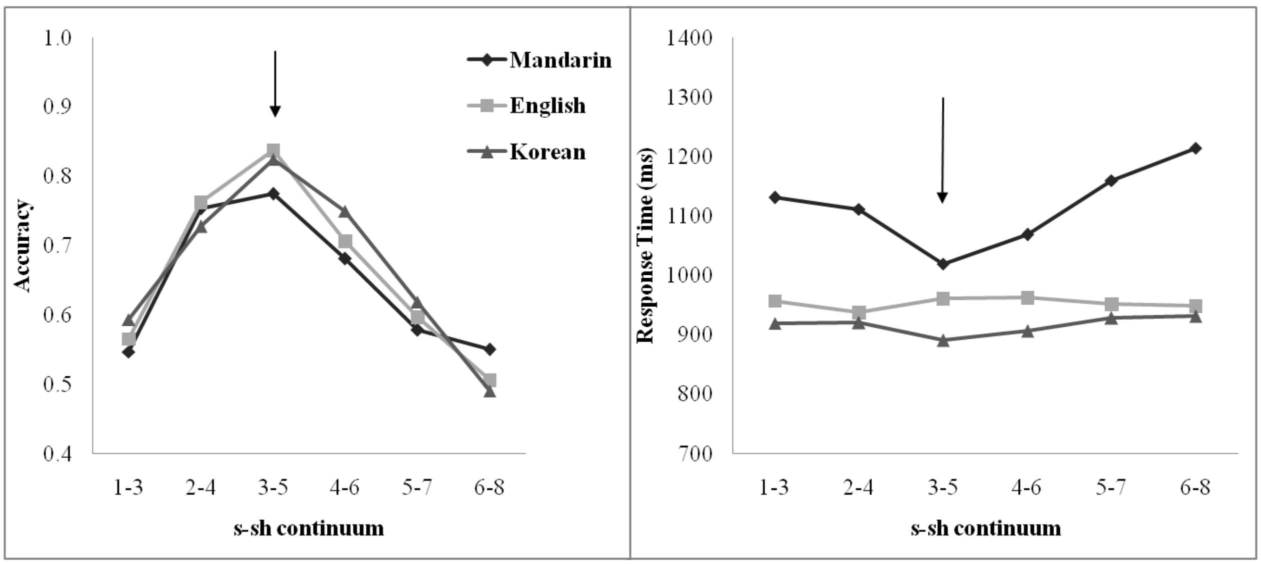
Figure 3-9 s-sh continuum response time



A follow-up repeated measures analysis showed a main effect of Language ($F(2,57)=4.114$, $p<.05$) suggesting that the RTs from the language groups were different. The condition Pair yielded a significant effect in the Mandarin group ($F(5,19)=2.641$, $p<.05$), but not in the other two groups (both $p>.5$), and this effect was best described by a quadratic trend ($F(1,19)=5.902$, $p<.05$).

We observe the same corresponding pattern for the Mandarin group in accuracy and response times, shown in Figure 3-10, as was seen for the *f-s* continuum. The valley of the Mandarin RT results in the right panel, indicated by the arrow, corresponded nicely to the peak of the accuracy on the left panel. In other words, the Mandarin listeners took less time and responded with higher accuracy between steps 3-5. We did not see this correspondence between accuracy and response times in the English and Korean groups.

Figure 3-10 s-sh continuum accuracy and response time results



The statistical results on response times are summarized in Table 3-8.

Table 3-8 Summary of discrimination response times results

***f-s* continuum**	***s-sh* continuum**
Language	*Language
***Pair	Pair
*Language x Pair	Language x Pair

The results for the Mandarin group seem to parallel the findings reported by other investigators that RT increases as a positive function of uncertainty (Pisoni & Tash 1974). We found shorter RTs when the two sounds were across a category boundary and longer RTs when the two sounds fell within a category that presumably pose more difficulty. The RTs of the English and Korean groups, however, did not show significant difference in terms of Pair. Note that the mean RTs of the English group and Korean group (English mean: 957 ms; Korean mean: 908 ms) were shorter than the Mandarin group overall (Mandarin mean: 1085 ms). The RT results seem to suggest that the English group and Korean group by-passed the phonological level of processing (faster response times and not patterning with the accuracy results), while the Mandarin group displayed a category-influenced judgment on the continuum (overall slower response times and patterning with the accuracy results). If the higher accuracy between steps 3-5 for the English group and Korean groups was a reflection of native language phoneme categories, we would have expected the same pattern for the RTs that we found for the Mandarin group.

A natural question to ask about the different RT patterns would be why only the Mandarin group should respond based on a phonological rather than an acoustic-phonetic level of processing. One possible explanation is that the stimuli were recorded by a Mandarin speaker. Therefore, the stimuli may have been more easily recognized as linguistic by the Mandarin speakers. Furthermore, Mandarin has been argued to have syllabic fricatives (Dong 1958; Chao 1968; Pulleyblank 1984; Ramsey 1987; Wiese 1997): structures that are sometimes analyzed as a syllable consisting of a fricative followed by an apical vowel ([tsɨ], [tsʰɨ], [sɨ], and [ʈʂʅ], [ʈʂʰʅ], [ʂʅ]) have alternatively been argued to be syllabic fricatives [ts̩], [ts̩ʰ], [s̩] and [ʈʂ̩], [ʈʂ̩ʰ], [ʂ̩]. This property of Mandarin Chinese might enable the Mandarin listeners to perceive the frications as complete syllables on the continua, and thus as speech sounds.

One might also wonder, if the task tapped into a low-level of processing for the English and Korean participants, why evidence of a category boundary should be present in the accuracy results. In other words, if the English and Korean participants were just listening to the acoustic differences between two members of the fricative pairs with *equal* acoustic distance, would we not expect a more gradient or non-categorical perception from the English and the Korean groups? However, previous studies have suggested that human beings tend to perceive sounds (even for non-speech sounds) categorically, rather than perceiving the small acoustic variations (Pisoni & Lazarus 1974; Pisoni 1977). The category boundary from the English and Korean results could just be a reflection of human's general auditory sensitivities.

3.5 Summary

The set of experiments discussed in this chapter compared the ability of speakers of English, Korean, and Mandarin to discriminate pairs from an eight-step continuum from *s* to *sh* and another continuum from *f* to *s* as comparison. The experiments were designed to determine whether the speakers from the three language groups showed a category boundary on the *s*-*sh* continuum. The accuracy results of this experiment did not yield a difference in the speakers' ability to discriminate the two sounds according to their language background. These results did not support either of the posited hypotheses and seemed to contradict the results of the first study, in which English and Mandarin speakers patterned together in contrast to Korean speakers.

Furthermore, the response time results showed that the English speakers patterned with the Korean speakers in that, overall, they took less time than did the Mandarin speakers in their discrimination of *s/sh*, and only the Mandarin speakers responded faster to pairs that they discriminated with higher accuracy. It was proposed that the experiment tapped into an acoustic-phonetic level of processing, instead of a phonological one to explained the discrepancy between the results of the similarity rating and discrimination experiments. The comparison of accuracy and RT patterns across the language groups suggests that only the Mandarin speakers were treating this as a linguistic task. The RTs of the English and Korean groups were shorter, possibly reflecting a low-level of processing, and did not vary according to Pair, whereas the Mandarin speakers took longer time in discriminating the stimuli, and the RTs varied according to Pair. The results suggest that this task did not tap into the listeners' native linguistic knowledge for the English and Korean groups due to the limited linguistic information provided. The longer RTs for the Mandarin group and the correspondence with accuracy suggest that the Mandarin listeners processed the stimuli at the phonological level.

Chapter 4 Semantic Priming

This chapter presents the results of an experiment investigating the priming relationship between *s* and *sh* in English, Korean, and Mandarin. Previous literature shows that variants of the same phoneme category participate in priming relationships while separate categories do not (Sumner & Samuel 2005; Ranbom & Connine 2007). In English, since *s* and *sh* are in contrast, we do not expect words containing *s* to prime words containing *sh*, and vice versa.[14] In both Korean and Mandarin, *s/sh* are in complementary distribution, but only in Korean do *s/sh* exhibit alternations. Examining the *s/sh* priming relationship in these two languages enable us to investigate the contribution of alternations vs. distribution in establishing sounds as members of a single phoneme category.

In Chapter 2 I showed that the Mandarin listeners patterned with the English speakers, rather than with the Korean speakers, in their similarity ratings for *s/sh*, which is expected if both the English and Mandarin speakers analyze these sounds as members of separate categories. However, the results of the discrimination experiment discussed in Chapter 3 were not consistent

[14] However, note that phonetically similar sounds might exhibit weak priming relationships (see Goldinger 1998 and the discussion in section 4.5).

with this pattern, since in terms of discrimination English, Mandarin, as well as Korean speakers showed a category boundary on the *s-sh* continuum. In Chapter 3 I suggested that the discrimination results may not be a good diagnostic for speakers' analysis of phonological relationships, since there results may have reflected acoustic rather than phonological processing. In this chapter, I will show that the results of the priming task also did not align with the similarity judgments, since in the priming experiment, the Mandarin speakers patterned with the Korean speakers. I will pursue possible explanations of this pattern.

The structure of this chapter is as follows. Section 4.1 summarizes previous relevant work in priming. Section 4.2 presents the methodology of the experiments. The results are presented separately for each language in section 4.3. Section 4.4 summarizes the chapter, with a discussion in section 4.5.

4.1 Introduction

It is well established that hearing a word primes the processing of an immediately following related word (i.e., semantic priming; e.g., Connine et al. 1993; Deelman & Connine 2001). Furthermore, previous research has found priming effects (i.e., faster reaction to experimental stimuli) between forms containing variant pronunciations of a category, but usually not between forms differing in sounds that belong to separate categories (Sumner & Samuel 2005; Ernestus & Baayen 2007; Ranbom & Connine 2007). For example, in a series of experiments using semantic priming and lexical decision tasks (i.e., how quickly people classify stimuli as words or non-words), Sumner & Samuel (2005) found that the target word *music* was primed by the semantically related word *flute*, when *flute* was articulated with any of the three variants of final [t]: canonical fully aspirated [t], coarticulated [ʔt̚] and glottalized [ʔ]. However, when the participants were presented with a contrastive phoneme /s/, instead of /t/, as in [flus], no facilitation was found in classifying the target word *music*. Note that the different variants of final [t] in Sumner and Samuel (2005) occur in free variation, so that listeners may be expected to have heard the three variants.

In contrast to the lack of priming relationship among [s] and [t] in English, Ernestus and Baayen (2007) have shown that even contrasting sounds may participate in a priming relationship when the two sounds participate in morpho-syntactic alternations. In Dutch, the voicing contrast in obstruents is neutralized in word-final position (/hand/ 'hand': [hant] *sg.*, [handə] *pl.*; /krant/ 'newspaper': [krant] *sg.*, [krantə] *pl.*), where voiced obstruents are pronounced as voiceless. In a series of lexical decision experiments, Ernestus and Baayen asked Dutch speakers to classify two types of forms with incorrectly voiced final obstruents as a word or non-word: (a) words with an underlying voiced final obstruent (e.g., *[hand] from /hand/); (b) words with an underlying voiceless final obstruent (e.g., *[krand] from /krant/). The results showed that the Dutch listeners classified type (a) forms as words more readily than type (b) forms, presumably because the lexical representations were activated when hearing the former type (e.g., /hand/) but not the latter (e.g., /krant/). The results suggest that even contrasting sounds (e.g., /d/ and /t/) may show a priming relation in a context in which the contrast is normally neutralized (e.g., word-final position).

To my knowledge no research has been done on the priming relationship of two sounds that do not contrast but that are in complementary distribution. This chapter describes an experiment investigating the priming relationships between *s* and *sh* in English (where the sounds are in contrast), Korean (where the sounds are in complementary distribution and participate in regular alternations), and Mandarin (where the sounds are in complementary distribution but do not alternate). Based on the Sumner & Samuel study showing lack of priming for contrasting sounds, we expect that forms differing in *s* vs. *sh* will not exhibit priming effects in English. The predictions for the other two languages are less clear. While the Sumner & Samuel study showed a priming relationship among the variant forms of final [t] in English, we cannot be certain whether this priming effect derives from the status of the [t] variants as members of the same category, or from the fact that these sounds are in free variation, which means that listeners have probably heard a single lexical item with varying pronunciations. Since in Mandarin and Korean listeners would never hear *s* and *sh* in the same phonological context, and therefore would never hear variant forms involving these two sounds, we cannot predict whether the two sounds should participate in a priming relationship in these two languages. However, the case of Korean vs. Mandarin still has the potential to shed light on the contribution of alternations vs. distributional evidence alone in establishing sounds as members of a single phoneme category. The crucial comparison for this point lies between Mandarin and Korean *s/sh*. If we find a *s/sh* priming relationship in Korean but not in Mandarin, this would presumably be attributable to the fact that Korean speakers have evidence for the *s/sh* relationship from both distribution and alternation while Mandarin speakers have evidence only from distribution, therefore supporting the *Distribution Plus Alternation* Hypothesis. Alternatively, if we find a priming relationship between the two sounds in both Korean and Mandarin, this would support the *Distribution Alone* Hypothesis.

4.2 Methodology

The priming experiment was designed to investigate the extent to which recognition of a target word was facilitated by a form derived by changing *s* to *sh*, or vice versa, in English, Korean, and Mandarin. The experiment employed a semantic priming paradigm using a lexical decision task. For all three languages, the facilitation effect of a *s/sh* change was compared with the priming effects when *s* and *sh* were changed to a contrasting sound. As the contrasting sound, *f* was chosen for Mandarin and English because of its contrast with *s/sh* in the two languages (e.g., minimal pair [**s**ǎ] 'spill' vs. [**f**ǎ] 'hair' in Mandarin, and ***s**ee*, ***f**ee*, and ***sh**e* in English). For Korean, where *f* does not occur, fortis *s'* and *sh'* were chosen since fortis vs. lax obstruents contrast in this language (e.g., on fricatives, [**s**al] 'flesh' and [**s'**al] 'uncooked grains of rice'; on stops, [**t**al] 'moon' and [**t'**al] 'daughter') (c.f., Table 1-3).

4.2.1 Participants

60 participants from each language group were recruited for this set of experiments. 40 of the participants had also participated in the previous two experiments, and the same recruiting requirements were followed for this experiment (see sections 2.2.1 and 3.2.1). The English group included 29 males and 31 females (aged 18-22); the Korean groups included 24 males and 36 females (aged 18-38) with an average self rating of 4.50 on their English ability; the Mandarin

group included 11 males and 49 females (aged 20-22) with an average self rating of 4.51 on their English ability. None reported any hearing deficiencies.

4.2.2 Norming pretest

A norming pretest was carried out to select semantically related word pairs for the experiment. A list of disyllabic *s* and *sh* ([ʃ/ɕ]) onset words was compiled in each of the languages (three lists in total). These lists were presented to 10 Taiwanese Mandarin speakers, 11 English speakers, and 5 Korean speakers, respectively. Participants were instructed to write down a related word for each item. The 36 most frequently reported semantic associates from the original lists (39% reporting rate for Mandarin, 45% for English, and 32% for Korean) were chosen as prime-target pairs for the *Related* condition. Another 36 primes were chosen for the *Unrelated* condition, and were matched up with unrelated targets (72 pairs in total). See Appendix B for the wordlists.

4.2.3 Designs and Materials

The stimuli were recorded in a sound-dampened room by a male native Mandarin phonetician, a male native English phonetician, and a female native Korean phonetician, respectively. Three versions of the stimuli corresponding to three experimental conditions (*Same*, *Swapping*, *Contrastive*) were created by splicing off the fricative onsets from the recorded stimuli using the Praat software package (Boersma 2001). The first version (the Same Condition) was created by splicing in a single token of *s* and *sh*, so that all the stimuli had physically the same *s* or *sh* token. (The same tokens of *s* and *sh* were used in creating the Swapping condition).

This method was used to avoid participants being influenced by the naturalness of the stimuli. The second version (the Swapping Condition) was created by cross-splicing the representative tokens of *s* and *sh* (e.g., [**ʃ**ambə] from ***s****amba*, [**s**ædoʊ] from ***sh****adow*), and the third version (the Contrastive Condition) was created by splicing in a contrastive onset, [f] in Mandarin and English (e.g., [**f**ambə] from ***s****amba*, [**f**ædoʊ] from ***sh****adow*), and fortis [s']/[ɕ'] in Korean. The intensity of all the stimuli was scaled to 65dB.

The design followed closely that of Sumner & Samuel (2005), which is a three-factorial design with two between-subject factors (Language and Condition) and one within-subject factor (Relation), as shown in Table 4-1.

Table 4-1   Semantic priming design

Between-subject factor	Language	→Mandarin, English, Korean
	Condition	→Same, Swapping, Contrastive
Within-subject factor	Relation	→Related, Unrelated
Dependent variable	Response times	

A sample wordlist is provided in Table 4-2. Sixty participants from each of the three language groups were randomly assigned to three experimental conditions (20 participants in each condition): *Same*, *Swapping*, and *Contrastive*.

Table 4-2 Sample wordlist

Language	Condition	Prime	Target
English	Same	[**s**ambə] ‘samba’	‘dance’
		[**ʃ**ædoʊ] ‘shadow’	‘darkness’
	c.f., Unrelated	[**s**æləd] ‘salad’	‘depress’
		[**ʃ**ʊgɚ] ‘sugar’	‘noise’
	Swapping	[**ʃ**ambə]	‘dance’
		[**s**ædoʊ]	‘darkness’
	c.f., Unrelated	[**ʃ**æləd]	‘depress’
		[**s**ʊgɚ]	‘noise’
	Contrastive	[**f**ambə]	‘dance’
		[**f**ædoʊ]	‘darkness’
	c.f., Unrelated	[**f**æləd]	‘depress’
		[**f**ʊgɚ]	‘noise’
Korean	Same	[**sɨ**k^{h}i] ‘ski’	[nun] ‘snow’
		[**ɕ**igje] ‘clock’	[ɕikan] ‘time’
	c.f., Unrelated	[**sɨ**ŋni] ‘victory’	[kip^{h}i] ‘depth’
		[**ɕ**ine] ‘city’	[kjeɾan] ‘egg’
	Swapping	*[**ɕɨ**k^{h}i]	[nun] ‘snow’
		*[**s**igje]	[ɕikan] ‘time’
	c.f., Unrelated	*[**ɕɨ**ŋni]	[kip^{h}i] ‘depth’
		*[**s**ine]	[kjeɾan] ‘egg’
	Contrastive	*[**ɕ’**ɨk^{h}i]	[nun] ‘snow’
		*[**s’**igje]	[ɕikan] ‘time’
	c.f., Unrelated	*[**ɕ’**ɨŋni]	[kip^{h}i] ‘depth’
		*[**s’**ine]	[kjeɾan] ‘egg’
Mandarin	Same	[**sɨ**-jaŋ] ‘breed’	[toŋ-wu] ‘animal’
		[**ɕi**-jan] ‘banquet’	[tɕje-hun] ‘wedding’
	c.f., Unrelated	[**sɨ**-tɕi] ‘driver’	[ts^{h}aɪkwa] ‘melon’
		[**ɕi**-jin] ‘attract’	[ʂən-min] ‘god’
	Swapping	*[**ɕɨ**-jaŋ]	[toŋ-wu] ‘animal’
		*[**si**-jan]	[tɕje-hun] ‘wedding’
	c.f., Unrelated	*[**ɕɨ**-tɕi]	[ts^{h}aɪkwa] ‘melon’
		*[**si**-jin]	[ʂən-min] ‘god’
	Contrastive	*[**fɨ**-jaŋ]	[toŋ-wu] ‘animal’
		*[**fi**-jan]	[tɕje-hun] ‘wedding’
	c.f., Unrelated	*[**fɨ**-tɕi]	[ts^{h}aɪkwa] ‘melon’
		*[**fi**-jin]	[ʂən-min] ‘god’

In the Same conditions, nothing was manipulated except for cross-splicing in a representative *s* or *sh* token (e.g., prime [**s**ambə] 'samba' and target *dance*, prime [**ʃ**ædoʊ] 'shadow' and target *darkness*). In the Swapping condition, *s* and *sh* were swapped (e.g., [**s**ambə]→[**ʃ**ambə], [**ʃ**ædoʊ]→[**s**ædoʊ]), and in the Contrastive condition, *s*/*sh* were changed to a contrastive sound (e.g., [**s**ambə]→[**f**ambə], [**ʃ**ædoʊ]→[**f**ædoʊ]). Note that the stimuli in the *Swapping* and *Contrastive* conditions in Mandarin and Korean created *illegal* sequences due to the distributional restrictions of the two languages: [s]/[s'] does not occur before the high-front vowel [i], and [ɕ]/[ɕ'] does not occur before the non-high-front vowel [ɨ]; [f] does not occur before [i]/[ɨ] in Mandarin. (The fact that the primes in the Mandarin and Korean stimuli contained illegal sound sequences distinguishes this study from the Sumner & Samuel study, and the possible implications of this difference will be discussed later.)

108 filler trials were added (18 with real word targets and 90 with pseudoword targets) to balance the word and non-word responses in the lexical decision task, and to avoid the development of strategies in the responses (e.g., the association of fricative-onset primes with word or non-word responses). The filler primes all had non-fricative onsets. For each list, there were 90 real word targets (72 targets semantically related to *s*/*sh* primes, and 18 filler targets), and 90 pseudowords, and in total 180 prime-target trials.

4.2.4 *Procedure*

Participants completed the experiment individually or in groups of up to four in separate booths, using a computer that was connected to a keyboard with keys labeled ‘yes’ and ‘no’ in their native language.[15] Each participant received one experimental list (the Same condition, Swapping condition, or Contrastive condition). Therefore, each participant heard each prime and target pair only once. All stimuli were presented binaurally over headphones at a comfortable listening level, and in a different random order for each participant, using E-Prime software (v2.0; Psychological Software Tools, Pittsburgh, PA). On each trial, participants were presented with an *auditory prime* (e.g., *samba*), followed by a 500 ms ISI, and followed by an *auditory target* (e.g., *dance*). Participants were presented with written instructions on the computer screen in their native language, and were asked to make a lexical decision (to judge if the target was a word) by pressing ‘yes’ or ‘no’ on the keyboard as soon as they were sure. The primes in which the onset fricatives were manipulated were not judged—only the targets were judged. Example stimuli were provided and each participant completed a practice session with 8 trials before the experiment. The experiment lasted approximately 12 minutes.

[15] The labels ‘yes’ and ‘no’ were put on the keys ‘d’ and ‘l’ on a keyboard because of their relative central position on the keyboard.

4.3 Results

We expect that English *s/sh* should not exhibit a priming relationship due to their status as contrasting phonemes. Therefore, for English speakers the response times for the lexical decision task should be the same for the Swapping and Contrastive conditions—that is, the response times for deciding if *dance* is a word should be the same after hearing [**ʃ**ambə] and after hearing [**f**ambə]—since the relation between *s* and *sh* is not different from that between *s/sh* and *f*. Hearing [**ʃ**ambə] or [**f**ambə] should not activate the word *samba* and thus the lexical decision of the target *dance* should not be facilitated.

Of particular interest here is the comparison between Mandarin and Korean response time patterns. We expect no priming relation between *s/sh* and a contrastive sound in either language (*f* in Mandarin and *sh'/s'* in Korean). Mandarin and Korean *s/sh* share the same distributional predictability; however, only Korean *s* and *sh* participate in alternations. If we find a *s/sh* priming relationship in Korean but not in Mandarin, manifested in the different response time patterns for the lexical decision for the Same and Swapping conditions, this would presumably suggest that the Korean speakers have analyzed these sounds as more closely related than the Mandarin speakers (*Distribution Plus Alternation Hypothesis*). On the other hand, if we see similar response time patterns for the lexical decision between the two language groups for the Same and Swapping conditions, then we conclude that both groups have arrived at similar analyses of the phonological relationship between the two sounds.

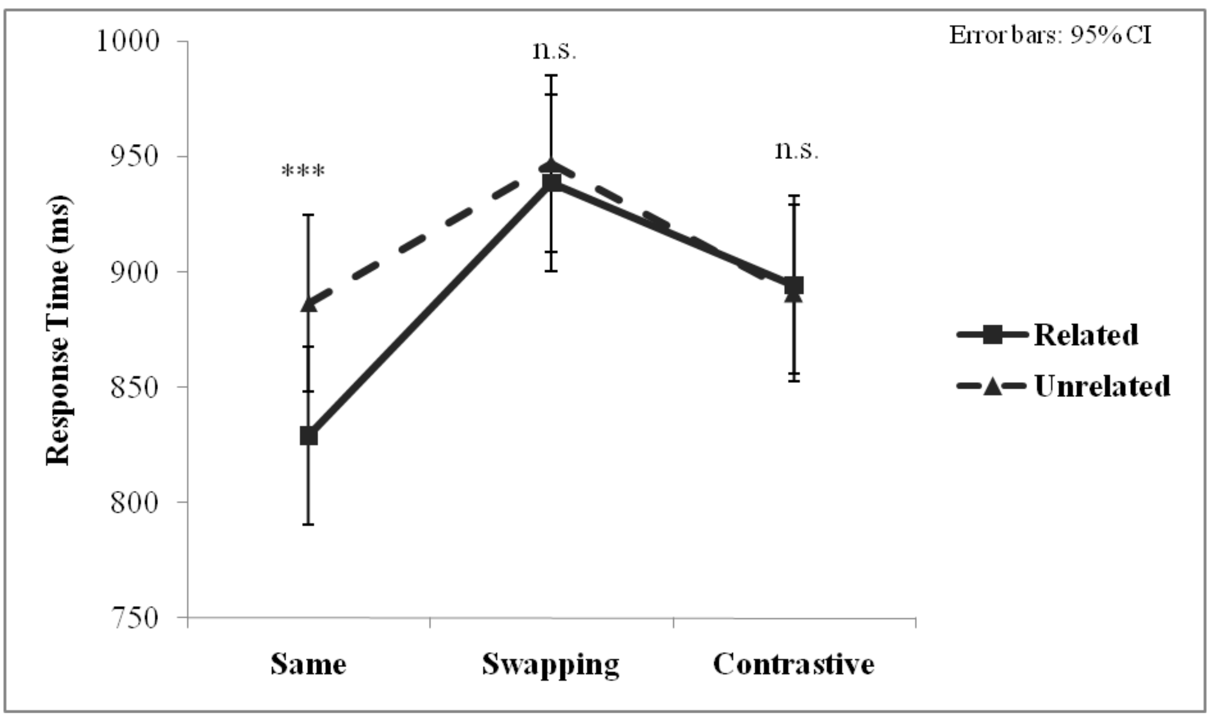
4.3.1 English

The mean response times and the priming effects (the difference between the Related and Unrelated conditions) for the three experimental conditions are shown in Table 4-3 with standard deviations in parentheses, and are illustrated in Figure 4-1.

Table 4-3 English lexical decision RT in ms and priming effect

Relation \ Condition	Same	Swapping	Contrastive
Related	829.05 (114.65)	938.82 (160.66)	894.47 (106.20)
Unrelated	886.65 (108.27)	947.11 (120.07)	890.72 (85.17)
Priming effect	57.6	8.29	-3.75

Figure 4-1 English lexical decision RT
(: p<.05; **: p<.01; ***: p<.001; n.s.: not significant)*



The response times for the unrelated prime-target pairs were taken as a baseline. Shorter response times observed in the related prime-target word pairs vs. the unrelated word pairs are taken to indicate a priming effect.[16] In Figure 4-1, the *x* axis represents the different conditions and the *y* axis represents the lexical decision times in milliseconds to semantically related targets (solid line/square markers) or to unrelated targets (dotted line/triangle markers). We can see from the response time differences between the related pairs and unrelated pairs in Table 4-3 and Figure 4-1 that only the *Same* condition yielded a significant difference, suggesting a priming effect when the onsets of the primes (*s* or *sh*) were identical (e.g., ***s****amba* or ***sh****adow*). In the other two conditions, in which the onsets were modified either by swapping *s* and *sh* (e.g., [**ʃ**ambə]), or changing *s/sh* to a contrastive sound (e.g., [**ʃ**ambə]), the distance between related pairs and unrelated pairs was very close, suggesting a lack of priming effect. This observation is confirmed by the statistical results.[17]

Two-way ANOVAs (**Condition** [Same, Swapping, Contrastive] **x Relation** [Related, Unrelated]) were performed across participants (*F*1) and items (*F*2). A main effect of Relation was found in the by-participant ANOVA ($F1(1, 57)=8.629$, $p<.05$), but not in the by-item

[16] The assumption of normality was met with the English lexical decision RTs. See Appendix E.5.
[17] Raw RT data were used in reporting the statistical results. Log transformed data yielded similar results to the raw data. For log transformed results, see Appendix D.

ANOVA ($F2(1, 210)=2.538$, $p=.113$).[18] A main effect of Condition was observed for the by-item ANOVA ($F2(2, 210)=14.453$, $p<.001$), but not for the by-participant ANOVA ($F1(2,57)=2.777$, $p=.071$). There was also a significant interaction for the by-participant ANOVA ($F1(2,57)=7.088$, $p<.05$, $F2(2,210)=2.085$, $p=.127$). Planned comparisons showed that targets preceded by related primes were identified more quickly than targets preceded by unrelated primes in the Same condition ($F1(1,19)=88.219$, $p<.001$, $F2(1,70)=9.116$, $p<.01$), but not in the other two conditions (Swapping $F1(1,19)=.240$, $p=.629$, $F2(1,70)=.128$, $p=.722$; Contrastive $F1(1,19)=.114$, $p=.740$, $F2(1,70)=.023$, $p=.880$).

The factor Condition yielded a significant effect in an analysis of the priming effects (cf., p.88 and Table 4-3), as illustrated in Figure 4-2 ($F(2, 117)=6.759$, $p<.05$). Pairwise comparisons showed that the main effect arose from the significant difference between the Same vs.

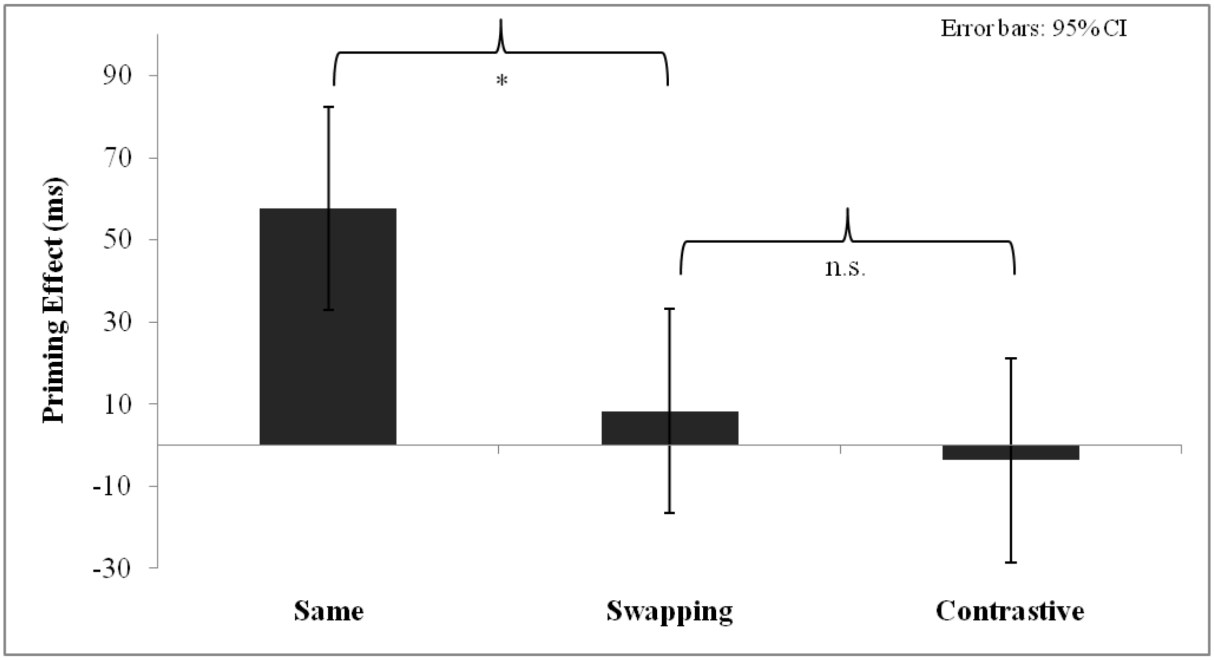
[18] The results reported here are based on the unscreened data. The screened data (excluding RTs when the lexical decisions were incorrect; 240 cases were tossed out), though producing lower standard deviations, yielded similar statistical results.

Relation \ Condition	**Same**	**Swapping**	**Contrastive**
Related	825.97 (111.42)	926.44 (134.10)	894.02 (106.12)
Unrelated	880.40 (111.10)	938.76 (115.77)	879.50 (83.80)
Priming effect	54.43	12.32	-14.52

With the screened data, there was a main effect of Relation ($F1(1, 57)=9.193$, $p<.005$, $F2(1,210)=4.722$, $p<.05$). Main effect of Condition was observed for the by-item ANOVA ($F2(2, 210)=9.967$, $p<.001$), but not for the by-participant ANOVA ($F1(2,57)=2.669$, $p=.078$). There was also a significant interaction for the by-participant ANOVA ($F1(2, 57)=12.210$, $p<.001$, $F2(2,210)=2.361$, $p=.097$). Planned comparisons showed that only in the Same condition, targets preceded by the related primes were identified more quickly than the unrelated primes ($F1(1,19)=61.571$, $p<.001$, $F2(1,70)=9.740$, $p<.005$), but not in the other two conditions (Swapping $F1(1,19)=.982$, $p=.334$, $F2(1,70)=.905$, $p=.345$; Contrastive $F1(1,19)=2.243$, $p=.151$, $F2(1,70)=.026$, $p=.872$).

Swapping conditions and the Same vs. Contrastive conditions (both $p<.05$). There was no significant difference between the Swapping & Contrastive conditions ($p=.775$).

Figure 4-2 English priming effect



The English results fit our predictions in that we found no facilitation in either the Swapping or the Contrastive condition, meaning that swapping *s* and *sh* was not different from changing *s/sh* to a contrastive sound *f*. The results showed a response time pattern consistent with the analysis of *s* and *sh* as contrastive phonemes in English. In other words, the relation between *s* and *sh* is not different from the relation between *s/sh* and a contrastive sound *f*. The statistical results are summarized in Table 4-4 (*: $p<.05$; **: $p<.01$; ***: $p<.001$).

Table 4-4 Summary of English semantic priming results

	***F*1 (by participant)**	***F*2 (by item)**
Omnibus ANOVA	Condition *Relation *Condition x Relation	***Condition Relation Condition x Relation
Same	***Relation	**Relation
Swapping	Relation	Relation
Contrastive	Relation	Relation

4.3.2 Korean

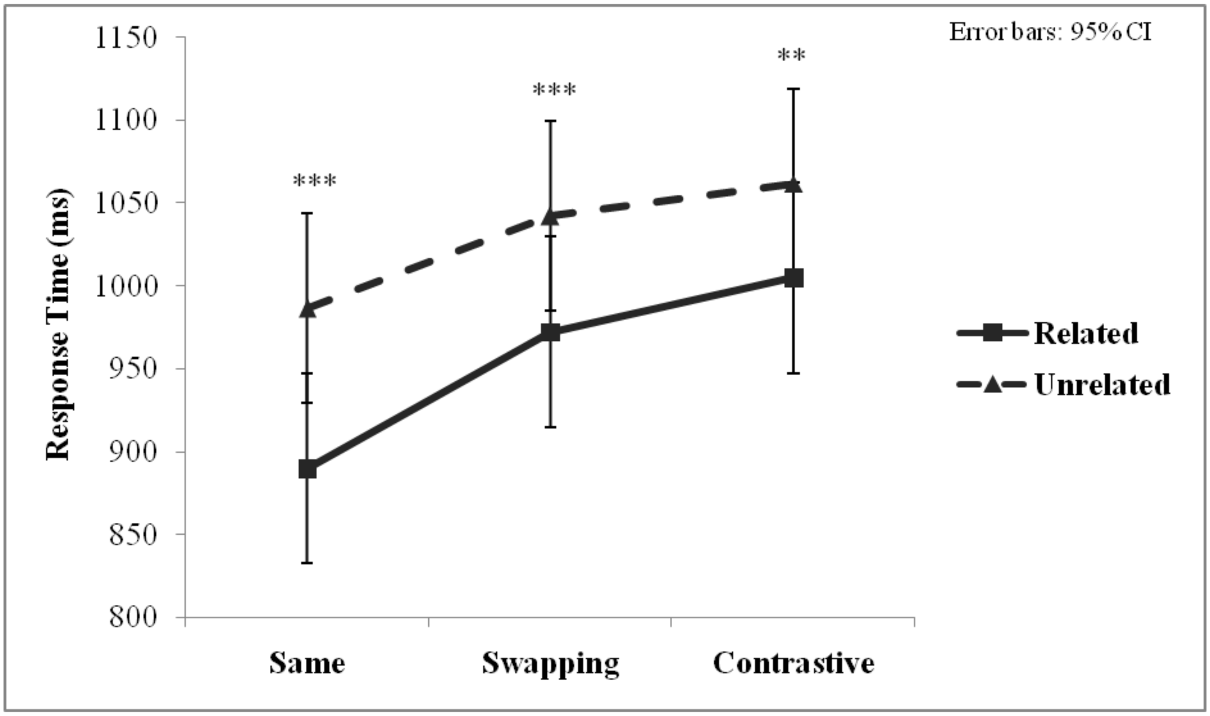
The mean response times and the priming effects (the difference between the related and unrelated prime-target pairs) for the three experimental conditions are shown in Table 4-5 with standard deviations in parentheses, and are illustrated in Figure 4-3.[19]

Table 4-5 Korean lexical decision RT in ms and priming effect

Relation \ Condition	**Same**	**Swapping**	**Contrastive**
Related	889.66 (74.66)	972.31 (156.40)	1004.86 (127.08)
Unrelated	986.69 (88.07)	1042.26 (172.40)	1061.95 (129.61)
Priming effect	97.03	69.95	58.09

[19] The assumption of normality was met with the Korean lexical decision RTs. See Appendix E.6.

Figure 4-3 Korean lexical decision RT



Two-way ANOVAs (**Condition** [Same, Swapping, Contrastive] **x Relation** [Related, Unrelated]) were performed across participants (*F*1) and items (*F*2). Overall, reaction times were significantly shorter for the related pairs than for the unrelated pairs ($F1(1, 57) = 88.895$, $p<.001$; $F2(1, 210) = 21.957$, $p<.001$). Planned comparisons showed that the targets preceded by the related primes were identified more quickly than the unrelated primes in all three conditions (Same $F1(1, 19)=65.342$, $p<.001$, $F2(1, 70)=15.558$, $p<.001$; Swapping $F1(1, 19)=25.896$, $p<.001$, $F2(1, 70)=5.973$, $p<.05$; Contrastive $F1(1, 19)=14.063$, $p<.005$, $F2(1, 70)=3.780$, $p=.056$). The factor Condition in the related prime-target pairs was significant ($F1(1, 57)=4.581$, $p<.05$, $F2(1, 105)=7.867$, $p<.05$). However, the factor Condition in the unrelated prime-target pairs was not significant for the by-participant analysis ($F1(2, 57)=1.684$, $p=.195$), but significant for the by-item analysis ($F2(1, 105)=4.851$, $p<.05$). [20]

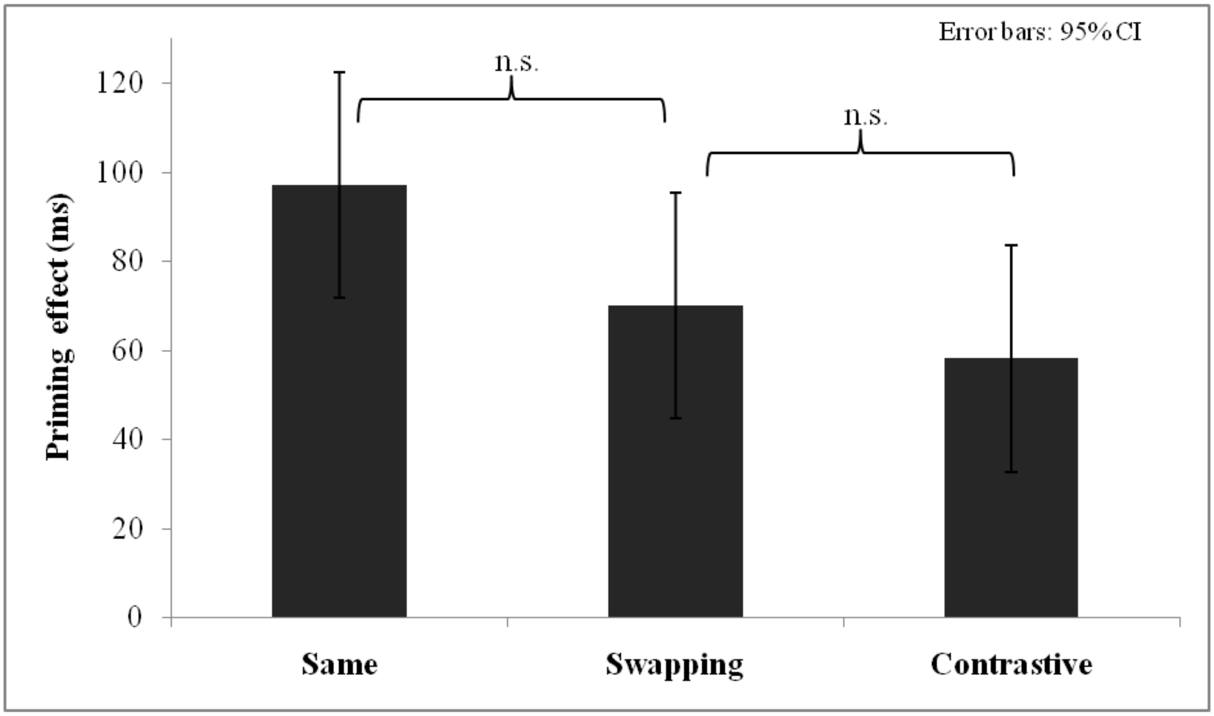
The factor Condition, however, did not yield a significant effect in an analysis of the priming effects ($F(2, 117)=2.531$, $p=.084$), meaning that the priming effects in the different conditions were not significantly different, as illustrated in Figure 4-4.

[20] The results reported here are based on the unscreened data. The screened data (excluding RTs when the lexical decisions were incorrect; 208 cases were tossed out), though produced slightly lower standard deviations, yielded the same statistical results.

Relation \ Condition	Same	Swapping	Contrastive
Related	886.47 (74.32)	964.88 (150.25)	997.64 (126.16)
Unrelated	973.18 (86.50)	1036.80 (185.71)	1055.32 (131.71)
Priming effect	85.11	71.92	57.68

With the screened data, reaction times were significantly faster for the related targets than for the unrelated targets ($F1(1, 57) = 83.36$, $p<.001$; $F2(1, 210) = 23.605$, $p<.001$). Planned comparisons showed that the targets preceded by the related primes were identified more quickly than the unrelated primes in all three conditions (Same $F1(1, 19)=53.021$, $p<.001$, $F2(1, 70)=15.408$, $p<.001$; Swapping $F1(1, 19)=28.19$, $p<.001$, $F2(1, 70)=7.905$, $p<.01$; Contrastive $F1(1, 19)=14.097$, $p<.05$, $F2(1, 70)=3.777$, $p=.056$). Simple effect of Condition in Related was not significant ($F1(2, 57)=4.448$, $p<.05$, $F2(2, 105)=8.611$, $p<.001$). However, the simple effect of Condition in Unrelated was not significant for the by-participant analysis ($F1(2, 57)=1.878$, $p=.162$), but significant for the by-item analysis ($F2(2, 105)=5.779$, $p<.05$).

Figure 4-4 Korean priming effect



In Korean, we did find a priming relationship between *s* and *sh* manifested by the significant priming effect between Related and Unrelated in the Same and Swapping conditions. In other words, swapping *s* and *sh* did not interfere with lexical retrieval. However, we also found a priming relationship between *s/sh* and the contrastive sounds *sh'/s'*, manifested by the significant priming effect in the Contrastive condition, and the strengths of the priming effects were not different across conditions. In other words, contrary to what we predicted, we found a priming relationship even among two contrastive sounds: changing *s/sh* to the contrastive sounds *sh'/s'* did not interfere with lexical retrieval.

The statistical results are summarized in Table 4-6.

Table 4-6 *Summary of Korean semantic priming results*

	***F*1 (by participant)**	***F*2 (by item)**
	Condition	***Condition
Omnibus ANOVA	***Relation	***Relation
	Condition x Relation	Condition x Relation
Same	***Relation	***Relation
Swapping	***Relation	*Relation
Contrastive	**Relation	Relation

4.3.3 Mandarin

The mean response times and the priming effects (the difference between the related and unrelated prime-target pairs) for the three experimental conditions are shown in Table 4-7 with standard deviations in parentheses, and are illustrated in Figure 4-5. [21]

[21] The results reported here are based on the unscreened data. The screened data (excluding RTs when the lexical decisions were incorrect; 253 cases were tossed out), though produced lower standard deviations, yielded the same statistical results.

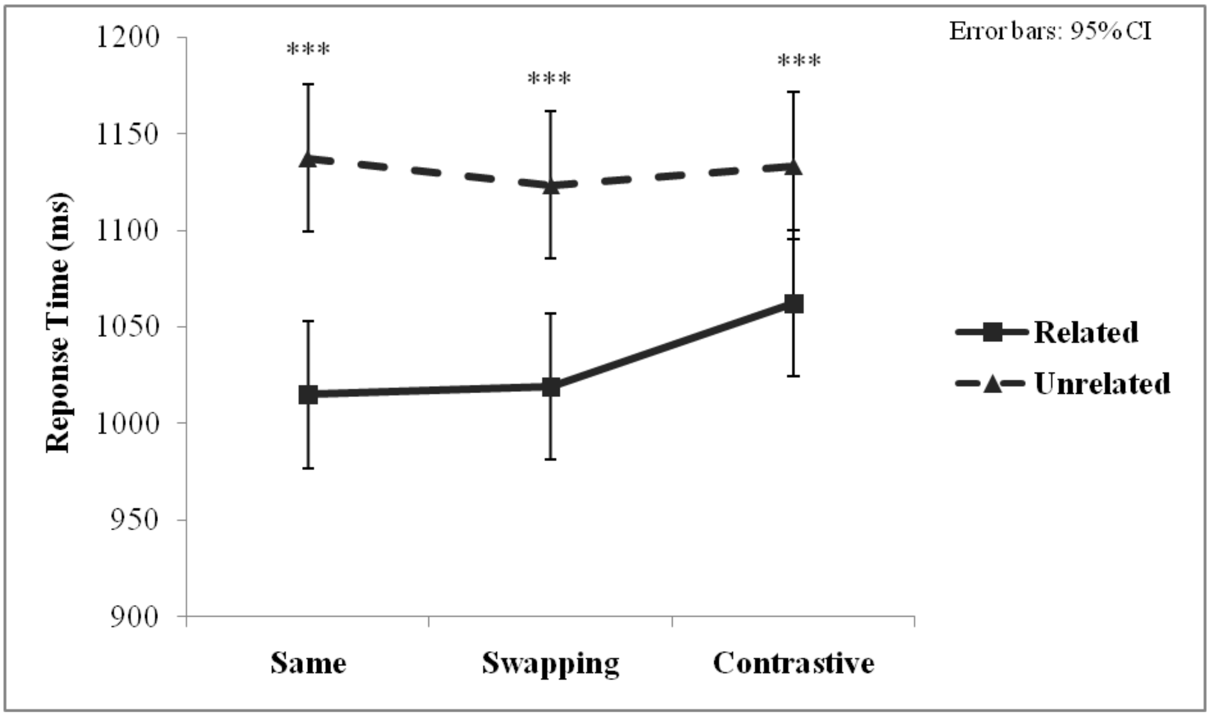
Condition / Relation	**Same**	**Swapping**	**Contrastive**
Related	1017.40 (113.72)	1011.64 (98.19)	1055.75 (127.16)
Unrelated	1126.17 (119.99)	1112.88 (105.07)	1119.63 (135.41)
Priming effect	108.77	101.24	63.88

With the screened data, reaction times were significantly faster for the related targets than for the unrelated targets ($F1(1, 57) = 125.173$, $p<.001$; $F2(1, 210) = 53.030$, $p<.001$). Planned comparisons showed that the targets preceded by the related primes were identified more quickly than the unrelated primes in all three conditions (Same $F1(1, 19)=58.144$, $p<.001$, $F2(1, 70)=22.802$, $p<.001$; Swapping $F1(1, 19)=32.224$, $p<.001$, $F2(1, 70)=24.762$, $p<.001$; Contrastive $F1(1, 19)=52.490$, $p<.001$, $F2(1, 70)=8.006$, $p<.01$). Simple effect of Condition in Related was not significant ($F1(2, 57)=.891$, $p=.416$, $F2(2, 105)=2.919$, $p=.058$) nor does the simple effect of Condition in Unrelated ($F1(2, 57)=.061$, $p=.941$, $F2(2, 105)=.123$, $p=.884$).

Table 4-7   Mandarin lexical decision RT in ms and priming effect

Relation \ Condition	Same	Swapping	Contrastive
Related	1014.94 (110.34)	1019.19 (106.21)	1062.26 (135.16)
Unrelated	1137.34 (116.27)	1123.48 (107.22)	1133.46 (139.57)
Priming effect	122.4	104.29	71.2

Figure 4-5   Mandarin lexical decision RT



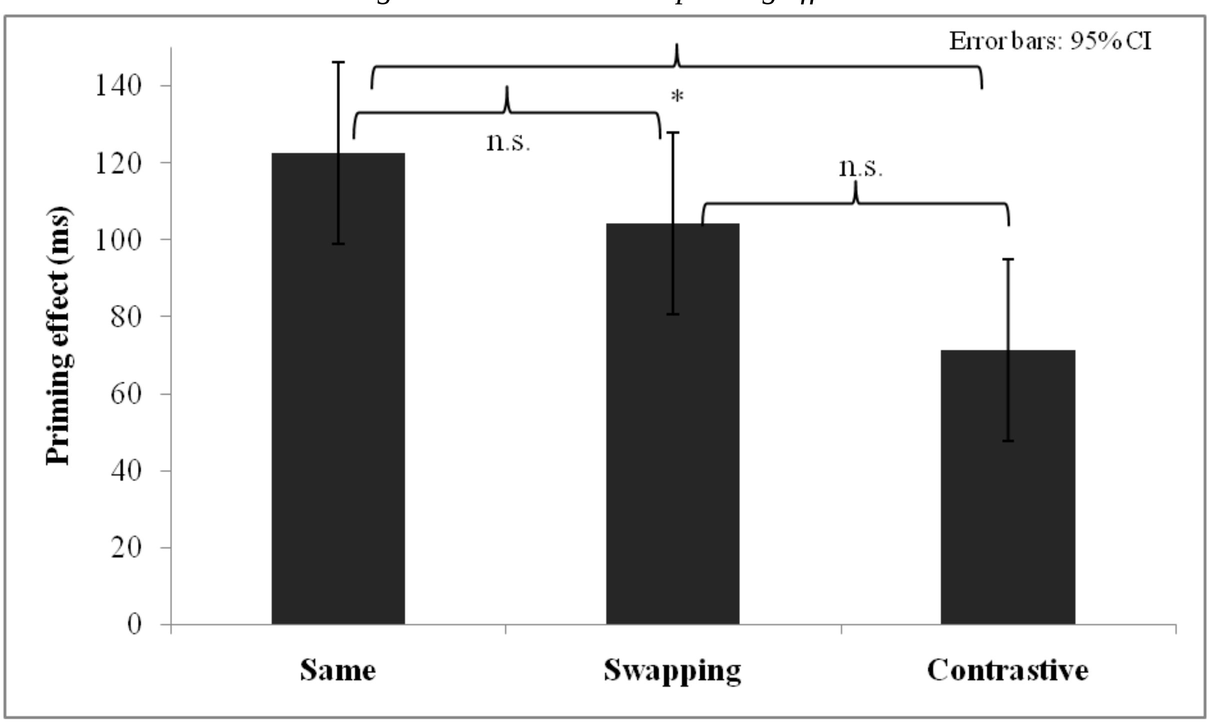
Two-way ANOVAs (**Condition** [Same, Swapping, Contrastive] **x Relation** [Related, Unrelated]) were performed across participants (*F*1) and items (*F*2).[22] Overall, the reaction times were significantly shorter for the related targets than for the unrelated targets (*F*1(1, 57) =

[22] The assumption of normality was met with the Mandarin lexical decision RTs. See Appendix E.4.

171.660, $p<.001$; $F2(1, 210) = 47.836$, $p<.001$). Planned comparisons showed that the targets preceded by the related primes were identified more quickly than the unrelated primes in all three conditions (Same $F1(1, 19)=85.445$, $p<.001$, $F2(1, 70)=23.178$, $p<.001$; Swapping $F1(1, 19)=44.004$, $p<.001$, $F2(1, 70)=19.767$, $p<.001$; Contrastive $F1(1, 19)=53.678$, $p<.001$, $F2(1, 70)=7.699$, $p<.01$). The factor Condition in the related prime-target pairs was not significant ($F1(2, 57)=.985$, $p=.380$, $F2(2, 105)=2.815$, $p=.064$), meaning that the response times of the related prime-target pairs in the three conditions were not significantly different. The response times for the unrelated prime-target pairs in the three conditions were not significantly different either ($F1(2, 57)=.069$, $p=.933$, $F2(2, 105)=.137$, $p=.873$).

The factor Condition yielded a significant effect in an analysis on the priming effects ($F(2, 117)=4.356$, $p<.05$), as illustrated in Figure 4-6.

Figure 4-6 Mandarin priming effect



Pairwise comparisons showed a significant difference in the priming effect only between the Same vs. Contrastive conditions ($p<.05$). The other two pairwise comparisons (Same vs. Swapping, and Swapping vs. Contrastive) were not significantly different (both $p>.1$).

Thus in Mandarin, as in Korean, we found priming relations between *s* and *sh*, as well as between *s/sh* and the contrastive sound *f*. Furthermore, no significant difference was found in priming effects between the three conditions. The statistical results are summarized in Table 4-8.

Table 4-8 Summary of Mandarin semantic priming

	***F*1 (by participant)**	***F*2 (by item)**
Omnibus ANOVA	Condition ***Relation *Condition x Relation	Condition ***Relation Condition x Relation
Same	***Relation	***Relation
Swapping	***Relation	***Relation
Contrastive	***Relation	**Relation

4.4 Summary

To summarize the results from the three language groups, we predicted that English *s/sh* should not participate in a priming relationship due to their status as contrasting phonemes, and that the response time for the lexical decision should be the same for the Swapping and Contrastive conditions. This effect was found in English, where only the Same condition yielded a significant effect of Relation, suggesting a priming effect when the onsets of the primes (*s* or *sh*) were identical (e.g., ***s**amba* primes *dance*, and ***sh**adow* primes *darkness*). In the other two conditions, in which the onsets were modified (e.g., Swapping: [**ʃ**ambə]; Contrastive: [**f**ambə]), no priming effect was found. In other words, *s* and *sh* were treated as separate categories by the English speakers, as were *s* or *sh* vs. a contrastive sound *f*.

Although we did not see priming effects in either the Swapping or Contrastive conditions, one might notice that the response times for the Swapping condition were longer than those for the Contrastive condition overall (942.97 ms vs. 892.6 ms; c.f., Figure 4-1). This could be due to the fact that in the Swapping condition, all the primes were non-words (e.g., [**ʃ**ambə] from *samba*; [**ʃ**ʌdeɪ] from *Sunday*), while the primes in the Contrastive condition were either words or non-words (e.g., non-word [**s**ɔldɚ] from *shoulder*, but actual word [**f**ɔldɚ] *folder*; non-word [**ʃ**ʌdeɪ] from *Sunday*, but actual word [**f**ʌdeɪ] *fun day*). It is possible that the slightly longer overall response times to the targets in the Swapping condition might reflect response when the primes were non-words. It might also be possible that the *word* primes in the Contrastive condition caused the slightly shorter overall response times to the target. To rule out this possibility, another ANOVA was run excluding the items that were words in the Contrastive condition (9 items out of 36 were excluded; see Appendix C for the wordlist), so that only the *non-word* items in the Swapping and Contrastive conditions were compared. The analysis yielded similar results. There was a main effect of Condition ($F2(1,183)=10.965$, $p<.001$). The factor Relation was significant only in the Same condition $F1(1,19)=76.483$, $p<.001$, $F2(1,61)=7.242$, $p<.01$), but not in the Swapping ($F1(1,19)=.418$, $p=.526$, $F2(1, 61)=.186$, $p=.668$) and the Contrastive conditions ($F1(1,19)=.489$, $p=.493$, $F2(1, 61)=.116$, $p=.735$).

One might also wonder if the overall longer response times for the Swapping condition than for the Contrastive condition came from the differences in length of the stimuli. Though the primes were different in length (*f* was shorter than *sh* [234 ms vs. 255 ms], length difference = 21 ms), participants were asked to judge the *targets* only. The targets were the same in all three experimental conditions and the response times were calculated from the beginning of each target. Thus, the length differences of the *primes* should not matter.

While the results from the English group were consistent with our predictions, for the Mandarin and Korean groups, on the other hand, we did not find a contrast in response times corresponding to same, allophonic, or contrastive sounds. Contrary to our predictions, we found a priming relationship both between *s/sh* and between these sounds and clearly contrastive sounds in both languages (the Contrastive conditions: *f* in Mandarin, and *sh'/s'* in Korean). Furthermore, we found a similar pattern—non-significant differences of priming effects among all three conditions—from both language groups. In other words, all three conditions exhibited priming effects, and the strengths of the priming were not different across conditions.

We now consider why we saw similar patterns between Mandarin and Korean groups in all three conditions, including the Contrastive condition. In the following section, I will pursue the possibility, from the results of a followup experiment, that the priming results actually reflect the effect of the distributional restriction, the fact that the swapping of *s/sh* and the changing of *s/sh* to the contrastive sounds in Mandarin and Korean (but not English) created forms containing sequences that are *illegal* in the native language. The hypothesis is that listeners actually misperceived these illegal sequences as legal ones, reflecting perceptual repair of illegal sequences at the level of lexical retrieval.

4.5 Discussion

As noted in section 4.2.3, the primes in the Swapping and Contrastive conditions were *illegal* sequences in both Korean and Mandarin: [s] ([s']) does not occur before the high-front vowel [i], and [ɕ] ([ɕ']) does not occur before the non-high-front vowel [ɨ]; [f] does not occur before [i] or [ɨ] in Mandarin. In English, the stimuli were all *legal* sequences: there are no phonotactic restrictions against *s*, *sh*, or *f* in a certain phonological environment (c.f., section 1.3.1). In making lexical decisions in a semantic priming experiment, the listeners hear a prime which is either semantically related or not related to the immediately following target, and they need to decide if the target is a word or not. When the prime is a real word, the corresponding word in the mental lexicon is presumably activated, and this activation facilitates the lexical decision of the semantically related target. When the prime is not a word, there should be no activation in the lexicon, and thus no facilitation of the immediately following word. However, the Korean and Mandarin participants heard primes with illegal sequences (*[si] and *[ɕɨ] in the Swapping condition, and *[fi] and *[fɨ] in the Contrastive condition). These differed from real words only in the substitution of a phonetically similar sound. Therefore, I hypothesize that in these cases, because the lexicon does not contain words starting with the illegal sequences, the phonetically closest corresponding real words were activated. In other words, the stimuli beginning with *[si] and *[ɕɨ] in the Swapping condition were ambiguous for the Korean and Mandarin listeners, in that the consonant [s] signals a following non-high-front vowel, but the vowel [i] signals a palatal onset. Similarly, the consonant [ɕ] signals a following high-front vowel, but the vowel [ɨ] signals a dental onset. Therefore, the Mandarin and Korean listeners were influenced by *top-down information from the lexical level*, and perceptually repaired the illegal sequences to the *phonetically closest* legal sequences that make *words*.[23] For example, the word [sɨ-jaŋ] 'breed' was activated when they heard the illegal *[ɕɨ-jaŋ] (in the Swapping condition), as well as when they heard *[fɨ-jaŋ] (in the Contrastive condition), since [s], [ɕ], and [f] are very similar phonetically. In English, where there are no phonotactic restrictions against *s* or *sh* in a certain phonological environment, on the other hand, the stimuli in the Swapping (e.g., [**ʃ**amba]) and Contrastive conditions (e.g., [**f**amba]) were legal. The hypothesis is that, upon hearing the legal non-words, the corresponding real words (e.g., *samba*) were not activated as readily because of the competing real words in the lexicon (e.g., ***sha****rp*, ***sha****lom*, ***sha****rk*, etc. for [**ʃ**amba], and ***fa****r*, ***fa****ther*, ***fo****llow*, ***fo****nd*, etc. for [**f**amba]) (Beckman & Pierrehumbert 2004).

The hypothesis that listeners perceived the illegal sequences as *phonetically similar* legal sequences that make *words* in lexical contexts is based on the previous speech processing and speech recognition literature, which shows that the *phonemic level* can be directly influenced by top-down information from the *lexical level* (e.g., Connine & Clifton 1987; Miller & Eimas 1995;

[23] Several speech recognition models have been proposed to explain the relationship between lexical processing and prelexical (e.g., phoneme) processing. For example, in TRACE, a model proposed by McClelland & Elman (1986), lexical influences result directly from lexical processes exerting top-down control over a prior process of phonemic analysis. In the Race model (Cutler & Norris 1979), phoneme identification can occur via a prelexical process or a lexical process, and the responses from speakers are basically the result of a race between these two processes. In the Merge model, "prelexical processing provides continuous information (in a strictly bottom-up fashion) to the lexical level, allowing activation of compatible lexical candidates. At the same time, this information is available for explicit phoneme decision making. The decision stage, however, also continuously accepts input from the lexical level and can merge the two sources of information" (Norris et al. 2000: 312). Note, however, that the discussion in this section is not committed to any particular model.

Wurm & Samuel 1997; Norris et al. 2000; Norris et al. 2003; Beckman & Pierrehumbert 2004; Hay et al. 2004), and the top-down mapping from the lexical level to the phonemic level can be dependent on *phonetic similarity* (e.g., Connine et al. 1993; Connine 1994; Connine et al. 1994; Connine et al. 1997). For example, Ganong (1980) conducted two identification experiments to test top-down lexical effects on phoneme categorization. A continuum of stops varying in VOT was used, where one endpoint was a word and the other a non-word (e.g., *task*-**dask*, **tash*-*dash*). The results showed that the boundary of the categorical perception was biased towards the endpoint which was a word. That is, in the context of *[t-d]ask*, English listeners were more likely to identify ambiguous stimuli as *task* than the non-word **dask*. On the other hand, in the context of *[t-d]ash* using the same VOT continuum, English listeners were more likely to identify ambiguous stimuli as *dash* than **tash*. In other words, when presented with the *same* continuum, listeners consistently shifted the phoneme boundary on the continuum in the direction of actual words. The results showed a top-down lexical effect on phoneme categorization: there is a tendency for listeners to make phonetic categorization that make *words*. The tendency is also shown to be greater when auditory information is ambiguous (in Ganong's case, the stimuli with sounds around the phoneme boundary on the continuum). As Norris et al. (2000) point out, when encountering incomplete or suboptimal perceptual input, the best the word recognition system can do is to identify the lexical representation in long-term memory that best matches the perceptual input.

Connine et al. (1993), in a series of semantic priming experiments using a lexical decision task, showed that phonetic similarity, even *across* phoneme boundaries, may also lead to priming. A lexical item could be activated by similar-sounding non-words that deviated in one or two linguistic features (e.g., voicing in ***p**attern* vs. ***b**attern*), while non-words that deviated by more than three linguistic features showed no priming effect (e.g., voicing, place of articulation, and manner in ***p**attern* vs. ***r**attern*). Along the same lines, Connine et al. (1997) investigated the relationship between the degree of lexical activation and phoneme perception, and showed that the detection of a phoneme in non-words was inhibited when the carrier non-word was similar to a real word (e.g., listeners had more difficulty detecting /p/ in a non-word**p**enefit* with a phonetically similar real word ***b**enefit* than in a word like *pulofit* with no phonetically similar counterpart real word). The lexical status of the real word *benefit* inhibited the perception of the phoneme /p/ when they heard ****p**enefit*, and the inhibition decreased as similarity of the carrier to a real word decreased (e.g., detecting *sh* in ****sh**enefit* is relatively easier). In other words, the listeners were biased by lexical information.

To come back to the Korean and Mandarin semantic priming results, we can assume that at the level of lexical retrieval, it is more difficult to detect the differences between phonetically similar sounds such as [s], [ɕ], and [f] when they appear in contexts where they are illegal. Therefore, it seems plausible that both the Mandarin and Korean listeners might have perceived the sounds [s], [ɕ], and [f] in the illegal sequences *[si] and *[ɕɨ] (in the Swapping condition) and *[fi] and *[fɨ] (in the Contrastive condition) as *phonetically similar* sounds that would be legal in that context, where the legal sequence would constitute a real word. As a result, the real words from which the illegal primes were modified were activated, facilitating the response to the semantically related targets. In English, however, though *s/sh/f* are phonetically similar, the stimuli in the Swapping (e.g., [**ʃ**amba]) and Contrastive conditions (e.g., [**f**amba]) were legal. Upon hearing the legal non-words, the corresponding real words (e.g., *samba*) were not activated as readily because of the competing real words in the lexicon, and thus no facilitation was found in the response times of the following targets.

4.5.1 Testing perceptual repair

To test the hypothesis that the Mandarin and Korean listeners perceived *illegal* sequences as *phonetically similar legal sequences* that activated the corresponding real words, I conducted a followup experiment. A fourth condition was added in Mandarin with a contrastive sound that is phonetically more distant from the previous experimental conditions (e.g., [t^h], the T condition). Half of the stimuli contained *illegal* sequences such as *[**t^{h}ɨ-**jaŋ], and the other half *legal* sequences such as [**t^h**i-jaŋ], as defined by the phonotactic restrictions of Mandarin.

The first prediction was that if listeners do perceptually repair the illegal sequences and map the stimuli to possible real words, we should see no or less priming in the T condition. The rationale behind this is that the phonetic deviation between the real words and primes in the T condition (that is, the deviation of *[**t^{h}ɨ-**jaŋ] from the real word [**sɨ-**jaŋ] 'breed') was greater than the deviation between the real words and primes in the Swapping condition (*[**ɕɨ-**jaŋ] from the real word [**sɨ-**jaŋ] 'breed') and the Contrastive condition (*[**fɨ-**jaŋ] from the real word [**sɨ-**jaŋ] 'breed'). The greater phonetic deviation of [t^h] from *s/sh/f* is based on the studies which show that stricture differences (e.g., [±continuant], [±sonorant], [±consonantal]) generate more perceptual dissimilarity than *place*—*s*, *sh*, and *f* deviate from one another in place, and *s/sh/f* deviate from [t^h] in continuancy (see Steriade (2008) and the references there).

The second prediction was that we should find priming effects or stronger priming effects only for the stimuli containing the *illegal* sequences (e.g., *[$\mathbf{t^{h}}$**ɨ**] in *[$\mathbf{t^{h}}$**ɨ-**jaŋ]), since these are the stimuli that did not have competitors in the lexicon. The effect of top-down lexical information is expected to be stronger in stimuli with the illegal sequence. The legal sequences [$\mathbf{t^{h}}$**i**] in [$\mathbf{t^{h}}$**i**-jan] would have other competitors in the lexicon (e.g., [$\mathbf{t^{h}}$**i**-hwan] 'change', [$\mathbf{t^{h}}$**i**] 'shave', [$\mathbf{t^{h}}$**i**-tʂ^{h}u] 'remove,' etc.), and thus the top-down lexical information is expected to be weaker. Facilitation in lexical decision is not expected in legal sequences. Examples for the fourth condition, along with the original design, are listed in Table 4-9.

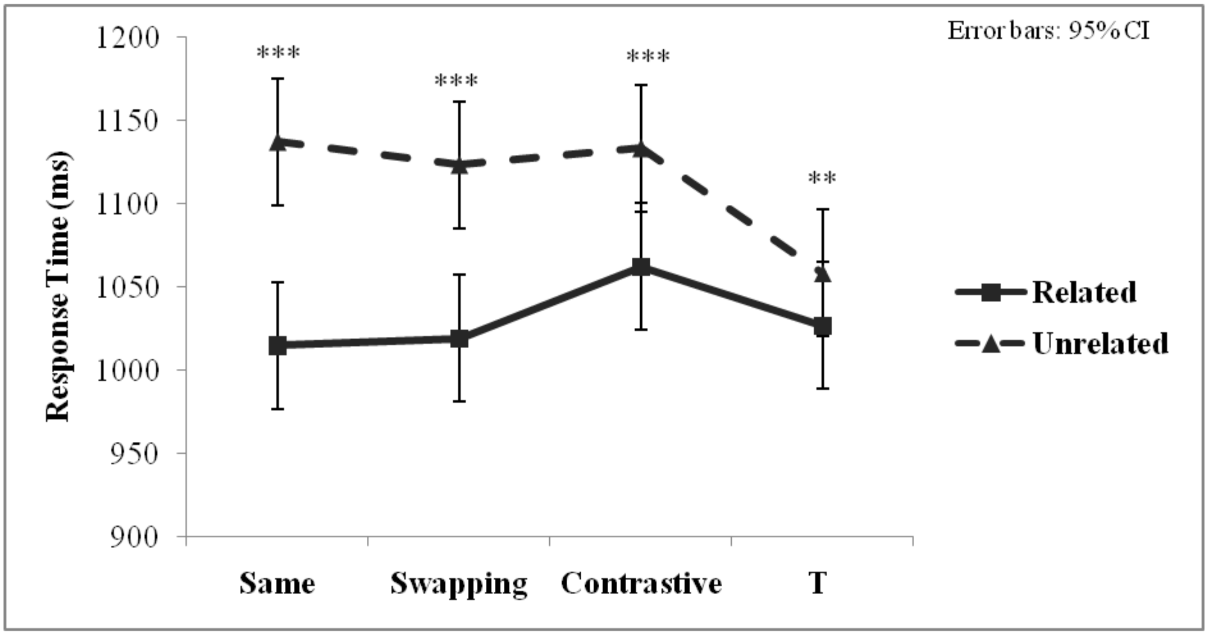
Table 4-9 Example wordlist with T condition

Mandarin	Same		[**s**ɨ-jaŋ] 'breed' [**ɕ**i-jan] 'banquet'	[toŋ-wu] 'animal' [tɕje-hun] 'wedding'
	Swapping		*[**ɕɨ-**jaŋ] *[**s**i-jan]	[toŋ-wu] 'animal' [tɕje-hun] 'wedding'
	Contrastive		*[**f**ɨ-jaŋ] *[**f**i-jan]	[toŋ-wu] 'animal' [tɕje-hun] 'wedding'
	T Condition	Illegal Legal	*[$\mathbf{t^{h}}$**ɨ-**jaŋ] [$\mathbf{t^{h}}$i-jan]	[toŋ-wu] 'animal' [tɕje-hun] 'wedding'

Another group of 20 Mandarin-speaking participants (10 male, 10 female, aged 18-42) were recruited for payment in Taiwan for this followup experiment. Their average self-rating of

English ability was 2.7 on a 7-point scale.[24] The same methodology as described in section 4.2 was followed. The results are illustrated in Figure 4-7.[25]

Figure 4-7 Mandarin T condition lexical decision RT



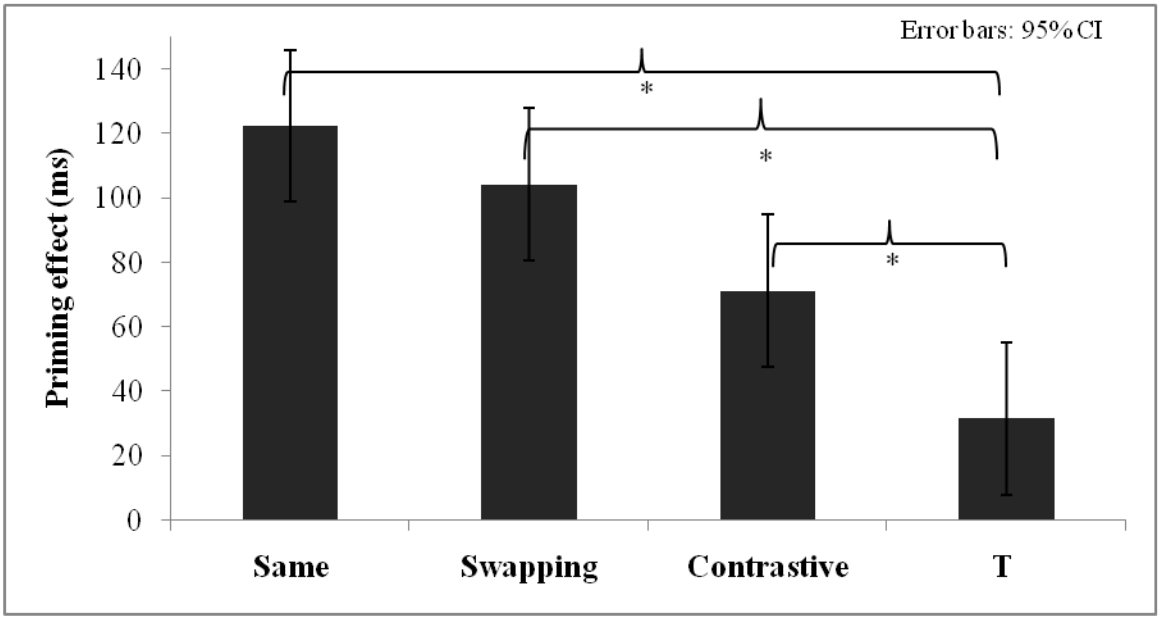
The factor Relation was significant in the T condition ($F(1,39) = 9.211$, $p<.005$). That is, the targets (e.g., [toŋ-wu] 'animal') following primes that were modified from semantically *related* words (e.g., [**sɨ**-jaŋ] 'breed'→*[**tʰɨ**-jaŋ]) were responded to faster than the targets following the primes that were modified from semantically *unrelated* words. In other words, there was still facilitation in retrieving the semantically related lexical items in the T condition.

[24] The participants in the follow-up study were recruited in a local church in Kaohsiung, Taiwan, not in a university. As a result, the self-rating of English ability was lower and the age range was wider.

[25] The assumption of normality was met with the Mandarin lexical decision RTs on the T condition. See Appendix E.4.

However, the priming effect was significantly smaller than in the other three conditions (Same vs. T, Swapping vs. T, Contrastive vs. T, all pairwise comparisons $p<.05$), as illustrated in Figure 4-8.

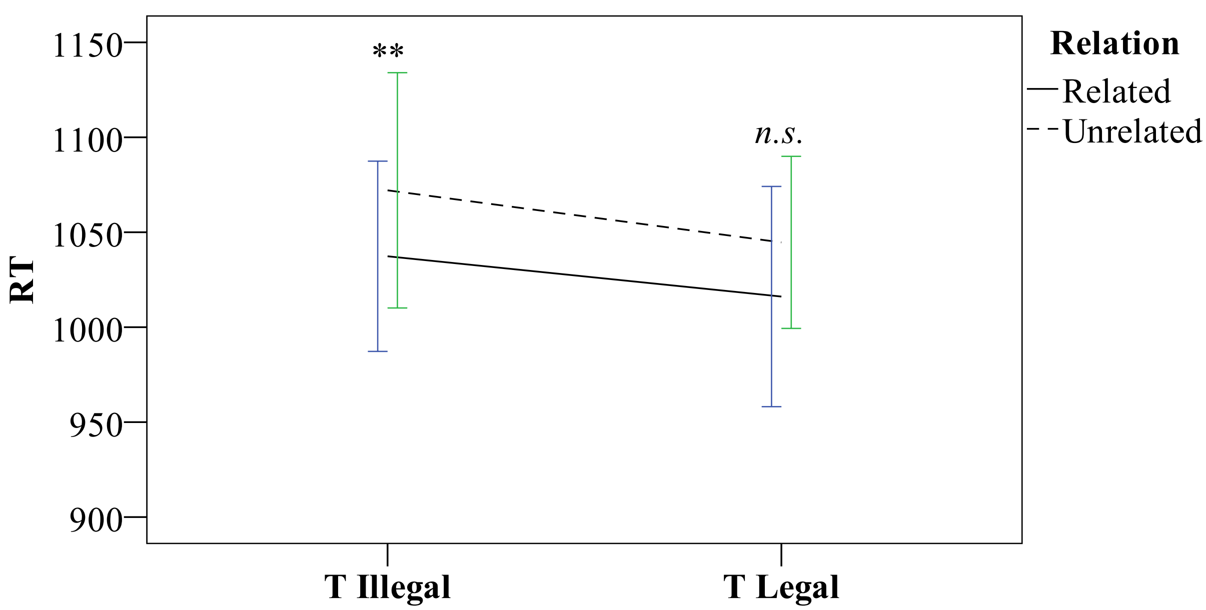
Figure 4-8 Mandarin T condition priming effect



The significant priming effect in the T condition suggests that the corresponding real words (e.g., [**sɨ**-jaŋ] 'breed') were still activated when the participants heard the modified words in the T condition (e.g., *[**t^hɨ**-jaŋ]). The activation of the real words in the lexicon facilitated the lexical decision on the following semantically related targets (e.g., [toŋ-wu] 'animal'), as we see from the faster response times for the related prime-target pairs (solid line) than for the unrelated prime-target pairs (dotted line). The priming effect in the T condition, however, was less than the other three conditions (Same, Swapping, and Contrastive). The smaller priming effect suggests that Mandarin listeners took longer to map the stimuli modified by substituting [t^h] to the corresponding real words than in the other three conditions (substituting *s/sh* in the Swapping condition, *f* in the Contrastive condition). The longer response times were presumably due to perceptual similarity: [t^h] deviates more from [s/ç/f].

In a post-hoc observation breaking the fourth condition into legal (T-legal) stimuli and illegal stimuli (T-illegal), as shown in Figure 4-9, although we did not see significant interaction of Legality and Relation ($F(1,19)=.014$, $p=.905$), we did see a significant effect of the factor Relation in the illegal sequences only (T-illegal, $F(1,19)=9.173$, $p<.01$). The forms containing legal sequences did not reach significance (T-legal, $F(1,19)=2.587$, $p=.124$). This suggests that the influence of the top-down lexical information was stronger in the stimuli with the illegal sequence, *[t^hɨ], since no competitors exist in the lexicon that contain the illegal sequence. The stimuli containing the legal sequence, [t^hi], however, had more competitors in the lexicon, and thus had weaker influence from the lexicon. In other words, the Mandarin listeners mapped the illegal stimuli more readily to the corresponding real words than the legal stimuli.

Figure 4-9 Mandarin T-legal and T-illegal lexical decision RT



To summarize the results of the followup experiment testing the hypothesis of perceptual repair, the T condition yielded a significant priming effect, and the priming was significantly smaller than the other three conditions. In addition, the Mandarin listeners responded faster to targets following *illegal* primes that were modified by changing [s/ɕ] of the real words to [t^h], but not to targets following *legal* primes. This suggests that the Mandarin listeners did perceptually repair the illegal sequences to *similar* legal sounding sequences that made *words*. In other words, the semantic priming experiments for the Mandarin and Korean groups, because of the illegal stimuli caused by the distributional restriction, involved a different level of processing from the English semantic priming.

4.5.2 Summary

To summarize this chapter, we found predicted results from the English group where only the Same condition yielded a significant effect of Relation, suggesting a priming effect when the onsets of the primes (*s* or *sh*) were identical (e.g., *samba* primes *dance* and *shadow* primes *darkness*), but no priming effect in the other two conditions, in which the onsets were modified (e.g., Swapping: [**ʃ**ambə]; Contrastive: [**f**ambə]). This supports the contrastive status of *s* and *sh* (and *f*) in English: no priming relation was found between *s* and *sh*, *s* and *f*, or *sh* and *f*. However, we did not see the expected results from the Korean group in that the Korean speakers showed a priming effect in all three conditions, and no evidence was found for the contrastive differences *s/sh* vs. *sh'/s'*. Mandarin speakers showed a similar lack of evidence of a difference in priming relations between *s* vs. *sh* and *s/sh* vs. *f*. The results are summarized in Table 4-10. ('*x*>*y*' represents ‘*x* has larger priming effect than *y*’;‘*x*≈*y*’ represents ‘the priming effect of *x* is not different from the priming effect of *y*.’)

Table 4-10 Summary of semantic priming results

English	Same > Swapping ≈ Contrastive
Korean	Same ≈ Swapping ≈ Contrastive
Mandarin	Same ≈ Swapping ≈ Contrastive

I then hypothesized that the similar pattern of the Mandarin and Korean speakers vs. the English speakers reflected the fact that in Mandarin and Korean (but not English) the altered primes contained illegal sequences, which I argued are perceived as phonetically similar real words. A followup experiment (T condition) was designed to test this hypothesis by changing *s/sh* to a contrastive sound, [t^{h}], that is phonetically more distant from the previous experimental conditions in Mandarin that produced *illegal* sequences for half of the stimuli (e.g., *[**t^{h}ɨ**-jaŋ]), and *legal* sequences for the other half (e.g., [**t^{h}**i-jan]). The results supported this hypothesis, showing that the T condition yielded a significant priming effect, but the priming was significantly smaller than the other three conditions (c.f., Figure 4-8), and that Mandarin listeners responded faster to targets following *illegal* primes that were modified by changing *s/sh* of the real words to [t^{h}], but not to targets following *legal* primes (c.f., Figure 4-9). The results support the hypothesis that the priming patterns in Mandarin and Korean resulted from a perceptual repair of illegal sequences at the level of lexical retrieval.

To come back to the initial rationale for the experiment, we expected the comparison of Korean vs. Mandarin to shed light on the contribution of alternations vs. distributional evidence alone in establishing sounds as members of a single phoneme category. However, because the use of illegal sequences (an inevitable outcome of substituting sounds in word contexts when these sounds are distributionally restricted) wiped out any possible effect of one-category vs. two-categories, the results do not provide evidence for the role of alternations vs. distribution.

The following chapter summarizes the results of the three experimental probes in the dissertation and concludes from the fact that the results did not yield a uniform pattern that there is no simple yes-or-no answer to decide that two sounds belong to a single category. Instead, the results suggest that the assignment of sounds to phoneme categories is gradient rather than absolute depending on both distribution and alternation. The implications of the results for different phonological theories will also be discussed in Chapter 5.

Chapter 5 Conclusion

This chapter begins with a review of the results of the experiments reported in previous chapters (section 5.1), followed by discussion of the implications of these results, first for the analysis of Mandarin fricatives and then for different phonological models (sections 5.2 and 5.3). I then discuss directions for future research (section 5.4).

5.1 Summary and conclusions

This dissertation began with the question of what sort of evidence causes native speakers to analyze two sounds as members of a single phoneme category. The goal was to investigate the contribution of two different types of evidence, *distribution* and *alternation*, to the processing of two coronal fricatives in three different languages: (i) English, in which the two sounds may occur in the same contexts, where the contrast between *s* and *sh* may signal differences in meaning (as in ***s**ee* vs. ***sh**e*), though the two sounds participate in limited morphological alternations; (ii) Korean, in which *s* and *sh* are in complementary distribution and participate in regular and productive morphological alternations; and finally (iii) Mandarin, in which *s* and *sh* are in complementary distribution but do not participate in allomorphic alternations. The different phonological relationships of the three languages are summarized in Table 5-1.

Table 5-1 Languages

	Mandarin	Korean	English
Distribution	√	√	
Alternation		√	(√)

The relationship between these two sounds in Mandarin has been a matter of controversy, with some researchers analyzing these sounds as independent phonemes (Cheng 1973; Yip 1996) and others analyzing them as variants of a single phoneme (Duanmu 2007; Wan 2010). Two hypotheses, the *Distribution Alone Hypothesis* and the *Distribution Plus Alternation Hypothesis*, were tested. If distributional predictability in the absence of alternations is a sufficient condition for causing speakers to group sounds into a single phonological category, as is the case for *s/sh* in Mandarin, we expect the results from the Mandarin group to be similar to those from the Korean speakers (*Distribution Alone Hypothesis*). If both distributional predictability and alternation are necessary to cause speakers to group sounds as a single category, then we expect the results from the Mandarin group and from the English group to pattern similarly (*Distribution Plus Alternation Hypothesis*). The predictions of the two hypotheses are summarized in Table 5-2.

Table 5-2 Predictions

Distribution Alone Hypothesis

	Mandarin	Korean	English
Distribution	√	√	
Alternation		√	(√)

Distribution Plus Alternation Hypothesis

	Mandarin	Korean	English
Distribution	√	√	
Alternation		√	(√)

Three previously established methods showing that variants of the same phoneme are processed differently than contrastive phonemes were used to probe the way in which speakers analyze the relationship between *s* and *sh*. These probes, however, did not yield consistent results. The experiment testing similarity ratings of *s* and *sh* for English, Korean, and Mandarin speakers was discussed in Chapter 2. Following previous work demonstrating that speakers tend to rate sounds that are allophonic variants in their language as more similar than sounds that are assigned to discrete phoneme categories, participants were asked to rate the similarity of *s* and *sh* ([ɕ] and [ʃ]), along with two other fricatives (*f*, *h*), embedded in three vowel contexts, [a_a], [i_i], and [u_u]. As expected, the Korean listeners, in whose language *s* and *sh* are in complementary distribution and participate in productive morphological alternations, rated these sounds as more similar than the English listeners did. The Mandarin listeners also rated the two sounds as significantly more different than did the Korean listeners. This suggests that the Mandarin speakers, with exposure only to distributional evidence, are less likely to group *s* and *sh* as members of the same category than Korean speakers, who are exposed to both distributional evidence and morphological alternation. Consequently, these results support the hypothesis that distribution alone is not sufficient to define phonological relationships (*Distribution Plus Alternation Hypothesis*). However, we also found that the similarity judgments for speakers of all three languages *varied* depending on the vowel context, even for the English speakers, in whose language *s* and *sh* may contrast in all the phonological environments included in the study. The effect of vowel context on similarity judgments suggests that there is not necessarily a simple answer to the question of whether speakers group two sounds as members of a single category or discrete categories.

The second experiment, reported in Chapter 3, tested English, Korean, and Mandarin speakers' discrimination of sounds on an eight-step synthesized *s*-*sh* continuum. The accuracy results of this experiment did not yield a difference in the speakers' ability to discriminate the two sounds according to their language background. Therefore, these results did not support either of the posited hypotheses. In fact, the response time results seemed to contradict the results of the first study, in which English and Mandarin speakers patterned together in contrast to Korean speakers. In the discrimination study, the response times of the English and Korean groups turned out to be shorter than those of the Mandarin group, and did not vary according to different fricative pairs, whereas the Mandarin speakers took longer in discriminating the stimuli, and showed different response times for different fricative pairs. I pursued the possibility that at least for some language groups, the experiment tapped into acoustic/phonetic rather than phonological processing, due to the limited linguistic information (i.e., frication only) provided, an interpretation that would explain the discrepancy between the results of the similarity rating and discrimination experiments. I proposed that only the Mandarin speakers were processing the stimuli as linguistic, which is reasonable given the fact that Mandarin has syllabic fricatives. English and Korean do not have syllabic fricatives, so it makes sense that the English and Korean speakers would process the stimuli as non-linguistic and respond only to acoustic differences.

The third experiment, discussed in Chapter 4, investigated the extent to which forms containing *s* primed forms containing *sh*, and vice versa. The results showed that for English speakers, *s* and *sh* did not show a priming relationship, parallel to the results for *s*/*sh* vs. other contrasting sounds. These results are consistent with the contrastive status of *s* and *sh* in English. However, for Korean and Mandarin speakers, a priming relationship was found not only between *s* and *sh* but also between these sounds and clearly contrasting sounds (*f* in Mandarin, and *sh'/s'* in Korean). I pursued the possibility that this unexpected effect resulted from the fact that some of the Mandarin and Korean forms contained illegal sequences, in which sounds were placed in vowel contexts where they do not normally occur (an inevitable outcome of substituting sounds in word contexts when these sounds are distributionally restricted). I suggested that speakers who heard the illegal sequences perceived them incorrectly, as containing the closest consonant that would be legal in that environment. Thus, the results did not provide evidence for the role of alternations vs. distribution, but instead reflected a perceptual repair at the level of lexical retrieval. Future research on the hypothesis of perceptual repair is suggested in section 5.4.

The results from the similarity rating experiment suggested that the evidence for assigning *s* and *sh* to a single category is stronger for Korean than for Mandarin speakers, supporting the *Distribution Plus Alternation Hypothesis*. Although the results from the discrimination and semantic priming experiments were inconsistent with the similarity ratings and in fact did not lend support to either of the hypotheses, these results may have reflected acoustic and lexical rather than phonological processing. Taken together, the results of these experiments further strengthen the view that different factors (distribution, alternation, phonological context, lexical context, etc.) may contribute to how listeners perceive the relationship between two sounds, and that phonological relationships may be gradient rather than categorical (Hall 2009).

The findings from the dissertation have several implications for the analysis of the Mandarin phoneme system and for models of the phonological grammar. I will first discuss the implications for the analysis of Mandarin palatal fricatives (section 5.2) and then the implications of these findings for different phonological models (section 5.3).

5.2 The analysis of Mandarin palatal fricatives

The controversy surrounding the analysis of Mandarin palatals results from the fact that the three palatals [ɕ, tɕ, $tɕ^h$] do not occur in the same contexts as three other series: the velars [x, k, k^h], the dentals [s, ts, ts^h], and the retroflexes [ʂ, tʂ, $tʂ^h$]. Under the assumption that as much redundant/predictable information should be removed from underlying representations as possible (following a principle of economy)—an assumption shared by many researchers in structuralist and generative phonology (Hockett 1942; Chomsky & Halle 1968; Clements 2003)—this complementary distribution "supplies pressure to eliminate the palatals as phonemes, and derive them from one of the other series" (Yip 1996: 770).

The assumption of economy is not uncontroversial (Prince & Smolensky 1993; Inkelas 1995; Yip 1996; Kager 1999; Krämer 2004). However, even accepting the assumption that palatals should be derived from one of the sequences with which they are in complementary distribution, the *lack of morphological alternations* makes any of the following analyses equally possible (c.f., Table 1-7 and the discussion in section 1.3.3).

Table 5-3 Analyses of Mandarin palatal fricatives

a. Surface palatals derived from underlying velars	
/x, k, k^{h}/ →[ɕ, tɕ, tɕ^{h}]	e.g., Chao (1934), Lin (1989), Chiang (1992), Wu (1994)
b. Surface palatals derived from underlying palatals	
/ɕ, tɕ, tɕ^{h} / →[ɕ, tɕ, tɕ^{h}]	e.g., Tung (1954), Cheng (1973), Yip (1996)
c. Surface palatals derived from underlying dentals	
/s, ts, ts^{h}/ →[ɕ, tɕ, tɕ^{h}]	e.g., Hartman (1944), Duanmu (2007)

Diachronically, there is a basis for deriving surface palatals from other sounds, since the palatals are residues of two historical processes, velar palatalization and dental sibilant palatalization (Dong 1958; Cheng 1973).[26] As a consequence, some researchers have argued, based on etymological evidence, that some palatals should be derived from underlying dentals, and others from underlying velars (Cheng 1968).

Turning to synchronic evidence, Chao (1934), along with other researchers (Lin 1989; Chiang 1992; Wu 1994), argues that the palatals [ɕ, tɕ, tɕ^{h}] should be identified with the velars /x, k, k^{h}/. The arguments come from palatal-velar alternations in two areas: word games and onomatopoeia. The word game normally infixes [ai.k] inside a syllable between the onset and the rhyme (e.g., [ma]→[m<u>ai.k</u>a]), as shown in (a-c) in Table 5-4. However, when the vowel of the original syllable is high, the infixed consonant is [tɕ] rather than [k], as in (d).

[26] The analysis of [ɕ, tɕ, tɕ^{h}] as underlying /ʂ, tʂ, tʂ^{h}/ has not been proposed in the literature presumably because the two series are etymologically unrelated (Chao 1934).

Table 5-4 [k]~[tɕ] alternation in May-ka language game (Chao 1931, 1934)

a.	ma	→	m<u>ai.</u>**k**a
b.	t^ha	→	t^h<u>ai.</u>**k**a
c.	k^hʊŋ	→	k^hw<u>ai.</u>**k**ʊŋ
d.	liŋ	→	lj<u>ai.</u>**tɕ**iŋ

The onomatopoeic expressions illustrated in Table 5-5 consist of reduplicated disyllables, where the first two syllables contain front vowel [i], and the last two syllables contain back vowel [u]. Crucially, the onsets of the first and third syllables are identical in (a), but in (b), [ɕ] appears before the front vowel and [k] before the back vowel.

Table 5-5 Onomatopoeia CV → Ci li Cu lu

a.	**t^h**i li **t^h**u lu	'slurping'
b.	**ɕ**i li **x**u lu	'eating fast'

These data showing alternations between the palatals and velars have been used to argue for palatals as underlying velars. However, Cheng (1973) argues that these patterns may be just a historical residue and should not affect synchronic phoneme categorization. Cheng thus concludes that palatals should be considered underlying segments because "although there are pieces of information favoring [palatals as underlying velars], there is no overwhelming evidence that I can find to support this view…. I have found no relation between the palatals and the other distributionally complementing series" (Cheng 1973: 40). Yip (1996) came to the same conclusion that the palatals should map faithfully to the underlying representation, based on the Optimality Theory notion of lexicon optimization—surface forms with the fewest violations of high-ranked constraints are assumed to faithfully map onto the underlying representation (Kager

1999)—that “learners will naturally internalize the forms closest to the surface, absent paradigm pressure [systematic morphological alternation] to do otherwise” (Yip 1996: 757).

Still another view is that the Mandarin palatals are derived from underlying dentals. Duanmu (2007) argues, on the basis of the distribution of glides, that [ɕ] is actually a surface realization of the consonant-glide combination (CG) /sj/ (/sa/→[sa] vs. /sja/→[ɕa]; /so/→[so] vs. /sjo/→[ɕo]). Duanmu further strengthens his argument that [ɕ] should be derived from underlying /sj/ by noting a variety of Mandarin Chinese in which the CG combination is pronounced as [s^j], instead of [ɕ].

Wan (2010) designed four experiments to investigate the psychological status of the palatals: onset similarity, sound contraction, sound similarity, and sound expansion. The experiments used a *misguide* method (Jaeger 1986) in which the participants were given sound sequences that are illegal (e.g., *[si], *[xi], *[ʂi]), and were told that the palatal series (the sounds that would create legal sequences in this context, e.g., [ɕi]) were missing. The participants then were asked to choose the perceptually closest sequence to [ɕi] from the three illegal sequences. The tasks were designed to determine which series—dentals, velars, or retroflexes—the participants identified most closely with the palatals. The results showed that participants favored the dentals in replacing the palatals significantly more often than the other series (the retroflexes and velars). Wan concluded from this asymmetrical response that the palatals should not be analyzed as independent underlying segments, and instead should be derived from the dentals. However, one can argue that the results only show that the palatals are perceptually more similar to the dentals than to the velars and retroflexes, but do not establish that palatals should be derived from underlying dentals. Furthermore, since the experiments employed stimuli with full syllables, one can also argue that the participants might have perceived the illegal sequences incorrectly, as argued for in discussion of the semantic priming results.

The results from the experiments in this dissertation provide new evidence bearing on this long-standing debate on the status of Mandarin palatal fricatives. Rather than comparing the similarity among the palatal fricatives and the other series within Mandarin, as in Wan's (2010) study, I compared Mandarin speakers' behavior with respect to *s* and *sh* with the behavior of speakers of other languages in which these sounds are in contrast or are closely related.

Because this dissertation compared the dental and palatal sounds only, the two series that are argued to be related in Duanmu (2007) and Wan (2010), we cannot rule out the possibility that Mandarin speakers identify the palatals with velars or retroflexes. It will be left for future research to carry out similar experiments with the other series of sounds (i.e., palatal-retroflex and palatal-velars) to see if the same results hold.

5.3 Implications for phonological relationships in phonological theory

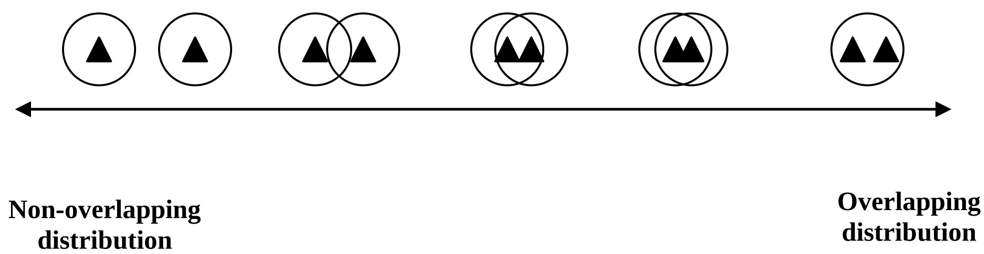
Phonological models differ in terms of their assumptions about sound relationships. Therefore, the findings in this dissertation also have implications for the assumptions and definitions made by different phonological models. Because the discussions of phonological relationships are abundant in the linguistic literature, a full treatment of this issue in different frameworks goes beyond the scope of this dissertation. I will focus on the following three questions:

1. Are sound relationships *categorical* or *gradient* (Goldsmith 1995; Hall 2009)? Must two sounds be either variants of the same category or of different categories, or can sound relationships fall somewhere on a *continuous scale* between contrast and allophony?
2. Do learners attempt to minimize the number of phonemes in lexical representations by deriving all *predictable* variants from a single phoneme? In other words, is the assumption of economy in the phoneme inventory desirable in a phonological analysis?
3. What type of evidence causes speakers to group sounds as members of the same or different category? Is distributional predictability sufficient to cause learners to assign two sounds to a single category, even in the absence of morphophonological alternations?

5.3.1 Categorical vs. Gradient

The finding that similarity ratings of *s/sh* from English, Korean, and Mandarin speakers varied depending on the vowel context and on the lack of consistent patterning within a language on the three experimental tasks suggests that there is not necessarily a simple answer to the question of whether speakers group two sounds as members of a single category or discrete categories. These findings cast doubt on approaches in which sound relationships are considered to be strictly categorical, supporting the position that phonological relationships may fall somewhere between contrast and allophony (e.g., Crowley 1998; Kristoffersen 2000; Moulton 2003; Ladd 2006; Rose & King 2007; Scobbie & Stuart-Smith 2008). In this camp, Hall (2009) argues, based on the degree of *predictability*, that "there are phonological relationships that are neither entirely predictable nor entirely unpredictable, but rather belong somewhere in between these two extremes" (Hall 2009: 307). Hall proposes that the phonological relationships of surface sounds fall on a continuum depending on the extent to which *the occurrence of a sound is predictable from its context*, as in Figure 5-1.

Figure 5-1 Varying degrees of predictability of distribution along a continuum (Hall 2009: 16)



Hall examines only the role of predictability from *distribution*, but acknowledges that "it is certainly not the case that distribution alone can accurately determine all phonological relationships. Nonetheless, in many cases, predictability of distribution is used as both a necessary and a sufficient condition for determining contrast and allophony" (Hall 2009: 11). Thus, although Hall's Probabilistic Model of Phonological Relationship (PPRM) assumes a notion of gradience that is drawn on a single dimension (predictability of distribution), she does not rule out the view of gradience proposed in this dissertation, in which multiple factors (e.g., distribution and alternation) may interact in determining the phonological relationships among sounds of a language.

Exemplar models (e.g., Goldinger 1996, 1997; Johnson 1997; Bybee 2001; Pierrehumbert 2001a, b; Bybee 2003; Pierrehumbert 2003a, b; Johnson 2005a, 2006; Pierrehumbert 2006a), also provide a natural account of gradience. In at least some exemplar models, all heard utterances are assumed to be stored, and grammar emerges as generalizations over these stored utterances. These generalizations are structured as a cognitive architecture in which fewer connections between two surface sounds signal contrast, and denser connections signal allophony. These gradient relationships in exemplar approaches can be defined by multiple dimensions in the cognitive structure minimally including generalizations over phonological environments, morphological paradigm, and phonetic similarity (Pierrehumbert 2003a). Thus, in exemplar models, as in PPRM, the relationships are presumably *gradient* as well, and gradience may be drawn from multiple dimensions.

5.3.2 Economy

If learners seek to minimize the number of phoneme contrasts in lexical representations, we would expect both Korean and Mandarin speakers to derive surface *s* and *sh* from a single underlying category. However, the results of the similarity ratings showed that the Mandarin speakers, like the English speakers, rated *s/sh*, two sounds with predictable distribution, as more different than did the Korean speakers, whose language provides evidence from both distribution and alternation for assigning these sounds to a single category. The results suggest that redundant information (e.g., the predictable distribution of Mandarin palatals) need not force sounds in complementary distribution to map onto the same underlying representation. These findings seem to challenge phonological approaches that assume *economy* in phoneme analysis, such as the traditional structuralist approach (e.g., Hockett 1942) and the SPE-type generative approach (*The Sound Pattern of English*; Chomsky & Halle 1968).

On the other hand, in output-based Optimality Theory (OT; Prince & Smolensky 1993), economy is generally assumed to play a much more limited role. Most researchers in OT assume no restrictions on the content of underlying representations, assuming *Richness of the Base*:

(2) *Richness of the Base:* no constraints hold at the level of underlying forms (Kager 1999: 19).

The outputs in an OT framework are evaluated by a set of ranked, violable constraints so that any input, even one containing illegal structures, will be mapped to a legal output, as defined by the constraint set. Thus in this approach, the predictable distribution of Mandarin palatals does not pose a problem at the level of underlying forms.

In approaches that abandon the notion of separate surface and lexical levels, the economy of underlying representations is irrelevant. Johnson (2005b, a) argues for exemplar-based generalizations over abstract category prototypes, on the basis that "if some sort of exemplar storage system is needed anyway, and if such a system can exhibit generalization behavior then why would one posit a parallel, totally redundant, prototype system?" (Johnson 2005b: 35). Though the concept of economy is irrelevant for exemplar-based models, the lack of an abstract level of underlying representations makes the implications of the results for this approach unclear.

5.3.3 Distribution vs. Alternation

The third question concerns the contributions of predictable distribution and morphological alternation, the two factors tested in this dissertation. In an Optimality Theory approach, in which the notion of Lexicon Optimization forces deriving two sounds from a single sound *only* when they are contained in a morpheme that *alternates*, the role of morphological alternation is made explicit (Inkelas 1995; Yip 1996) while distributional predictability in grouping sounds as the same underlying representation is diminished. The results of the similarity rating experiment, in which the Mandarin speakers rated *s/sh* as more different than did the Korean speakers, are consistent with the view that the additional evidence from alternation that the Korean speakers are exposed to had an effect. These findings suggest that alternation, as well as distribution, affects the grouping of *s/sh*, supporting models in which multiple factors (e.g., distribution, morphological alternation, and phonetic similarity) may all contribute to the formation of sound categories.

To sum up the discussion above, the findings in this dissertation are consistent with models in which phonological relationships are *gradient* rather than categorical, in which economy does not force the exclusion of redundant information in underlying representation, and in which multiple factors may contribute to the formation of phoneme categories.

5.4 Future Research

Several important questions have emerged from this dissertation. To conclude, I will outline some additional questions and three areas for future research to address these questions: the effect of orthography, the weighting of different factors determining phonological relationships, and the 'perceptual repair' hypothesis suggested in the priming experiment.

The first question involves the role of orthography in the perception of *s* and *sh* in these three languages. As mentioned in 3.2, Korean *s* and *sh* are represented with the same orthographic symbol 'ㅅ', which is pronounced as *sh* when the following vowel is a high-front vowel/glide, and *s* elsewhere. On the other hand, English *s* and *sh* are generally represented with different spellings (e.g., ***s**ee*, ***sh**e*) though sometimes with the same or similar spelling (e.g., ***s**ure* [ʃ]). In Mandarin, the standard writing system is non-alphabetic. However, a phonetic system is used as well, as is taught to school-age children (Zhuyin Fuhao/Bopomofo in Taiwan, and Hanyu Pinyin in China). In this system, *s* and *sh* are represented with different phonetic symbols: *s* is represented with 'ㄙ' in Zhuyin Fuhao and 's' in Hanyu Pinyin, and *sh* is represented with 'ㄒ' in Zhuyin and 'x' in Pinyin. School-age children learn the phonetic systems before the ideographic system. To determine whether the orthography affected the results of the experiments, one area for future study is to see if the same results hold for pre-school age children.

The second question concerns the weighting of different factors determining phonological relationships. To what extent can alternation, independent from other criteria (e.g., predictability of distribution or phonetic similarity) lead learners to group sounds together? Is there an inherent bias, or is the weighting of different factors derived from different degrees of exposure? This question could be explored by using artificial language learning experiments in which different factors are manipulated. This direction of research also has implications for theories of language change, specifically for explaining phoneme merging or splitting in the history of a phoneme system. If learners are inherently biased toward using certain criteria in grouping sounds as the same category, then those criteria are expected to enhance the merging of different phonemes.

A third direction for future research involves testing the hypothesis that in the priming experiment, Mandarin and Korean speakers heard illegal sequences as legal ones, due to a 'perceptual repair' of the illegal stimuli at the level of lexical retrieval. Since the hypothesis is that the repair happens at the lexical level, we expect to see different processing of the same illegal sequence when it is embedded in lexical contexts. A well-established component, Mismatch Negativity (MMN), in the Event Related Potential (ERP) literature can be pursued to look into the hypothesis of perceptual repair at the lexical level. MMN is a negative-going shift of the electrical brain activity (electroencephalogram, or EEG) upon 100 ms to 250 ms after a (discriminable) change in acoustic stimuli that can be elicited even in the absence of attention (e.g., Dehaene-Lambertz 1997; Näätänen et al. 1997; Näätänen 2001; Eulitz & Lahiri 2004; Luck 2005; Näätänen et al. 2007).[27] Steinberg et al. (2009, 2010) investigated the effects of language-specific phonotactic restrictions on pre-attentive auditory speech processing by native speakers of German. In German, a vowel and a following dorsal fricative (the palatal [ç] and velar [x]) agree in their phonological specifications for tongue backness; in other words, the palatal fricative [ç] can only occur after a front vowel, and the velar [x] can only occur after a back

[27] The discriminable auditory change could be a simple sound such as a sinusoidal tone, or a complex sound such as a phoneme, or a complex spectrotemporal pattern (Näätänen 2001).

vowel. Steinberg et al. (2010) looked at two comparisons, *[ɔç] and [ɛç] vs. [ɔʃ] and [ɛʃ], in which the first comparison contained a phonotactically illegal syllable. The results showed that a MMN corresponding to the vowel changes (around 100 ms to 120 ms) was observed for both comparisons; however, an additional MMN occurred approximately 350 ms after the onset of the stimuli was found only for the first comparison that contained the phonotactically illegal syllable. Steinberg et al. suggest that this additional MMN reflects the phonotactic ill-formedness of *[ɔç]. Along the same lines, we would expect the comparison of *[si] and [ɕi] to elicit a MMN corresponding to the consonant change as well as the additional MMN corresponding to the ill-formedness of *[si] in Mandarin speakers' pre-attentive auditory speech processing. If the perceptual repair occurs at the level of lexical retrieval—that is, if Mandarin speakers hear the illegal sequences incorrectly as the phonetically closest legal sequences—we would expect to see this additional MMN go away when the illegal sequence [si] is embedded in word contexts (e.g., comparison of *[**si**-jan] and [**ɕi**-jan] 'breed').

To conclude, the findings of the dissertation suggest that there is not a simple answer to whether speakers group two sounds as members of a single category or discrete categories, that phonological relationships may be gradient rather than categorical, and that the two factors tested, distribution and alternation, may both contribute to how listeners perceive the relationship between two sounds. Much work has to be done to tease apart the relative contributions of different factors. The dissertation suggested future directions in this respect.

References

Anderson, Stephen R. (1985). *Phonology in the Twentieth Century: Theories of Rules and Theories of Representations.* Chicago: The University of Chicago Press.

Babel, Molly & Keith Johnson. (2010). Accessing psycho-acoustic perception with speech sounds. *Laboratory Phonology* 1: 179-205.

Banksira, Deqif Petros. (2000). *Sound Mutations: The Morphophonology of Chaha.* Philadelphia: John Benjamins Publishing Co.

Baudouin de Courtenay, Jan. (1972). An attempt at a theory of phonetic alternations. *A Baudouin de Courtenay anthology: the beginnings of structural linguistics*, p.144-212. Indiana University Press.

Beckman, Mary E. & Janet Pierrehumbert. (2004). Interpreting 'phonetic interpretation' over the lexicon. *Papers in Laboratory Phonology VI,* p.13-38. Cambridge University Press.

Beckman, Mary E. & Janet B. Pierrehumbert. (2000). Positions, probabilities, and levels of categorization. *Keynote address, Eighth Australian International Conference on Speech Science and Technology*, Canberra, Dec. 4-7, 2000.

Best, Catherine T., Gerald W. McRoberts & Nomathemba M. Sithole. (1988). Examination of perceptual reorganization for nonnative speech contrasts: Zulu clicks discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance* 14: 345-60.

Bloch, Bernard. (1948). A set of postulates for phonemic analysis. *Language* 24: 3-46.

Bloch, Bernard. (1950). Studies in colloquial Japanese IV phonemics. *Language* 26: 86-125.

Boersma, Paul. (2001). Praat, a system for doing phonetics by computer. *Glot International* 5:9/10: 341-45.

Booij, Geert. (1999). *The Phonology of Dutch.* Oxford University Press.

Boomershine, Amanda, Kathleen Currie Hall, Elizabeth Hume & Keith Johnson. (2008). The impact of allophony versus contrast on speech perception. *Contrast in Phonology*, (eds) Avery, Peter, Elan Dresher & Keren Rice, p.143-72. Berlin: Mouton de Gruyter.

Bullock, Barbara E. & Chip Gerfen. (2005). The preservation of schwa in the converging phonological system of Frenchville (PA) French. *Bilingualism: Language and Cognition* 8: 117-30.

Bybee, Joan. (2001). *Phonology and Langauge Use.* Cambridge University Press.

Bybee, Joan. (2003). Mechanisms of change in grammaticization: The role of frequency. *Handbook of Historical Linguistics*, (eds) Joseph, Brian & Richard Janda, p.602-23. Oxford: Blackwell.

Chao, Yuen-Ren. (1931). Fan-qie yu ba zhong [Eight varieties of secret language based on the principle of fanqie]. *Bulletin of the Institute of History and Philology, Academia Sinica* II: 320-54.

Chao, Yuen-Ren. (1934). Tthe non-uniqueness of phonemic solutions of phonetic systems. *Bulletin of the Institute of History and Philology* 4: 363-97.

Chao, Yuen-Ren. (1968). *A Grammar of Spoken Chinese.* Berkeley and Los Angeles: University of California Press.

Cheng, Chin-Chuan. (1968). *Mandarin phonology*. University of Illinois at Urbana-Champaign.

Cheng, Chin-Chuan. (1973). *A Synchronic Phonology of Mandarin Chinese.* De Gruyter Mouton.

Chiang, Wen-Yu. (1992). *Prosodic Phonology and Morphology of Affixation in Chinese*. University of Delaware.

Chomsky, Noam & Morris Halle. (1968). *The Sound Pattern of English.* Cambridge, Massachusetts, London, England: The MIT Press.

Clements, G. N. (2003). Feature economy in sound systems. *Phonology* 20: 287-333.

Connine, Cynthia M. (1994). Vertical and horizontal similarity in spoken-word recognition. *Perspectives on Sentence Processing*, (eds) Clifton, Charles, Lyn Frazier & Keith Rayner, p.107-20. Hillsdale: Erlbaum.

Connine, Cynthia M., Dawn G. Blasko & Debra Titione. (1993). Do the beginnings of spoken words have a special status in auditory word recognition? *Journal of Memory and Language* 32: 193-210.

Connine, Cynthia M., Dawn G. Blasko & Jian Wang. (1994). Vertical similarity in spoken word recognition: Multiple lexical activation, individual differences, and the role of sentence context. *Perception & Psychophysics* 56: 624-36.

Connine, Cynthia M. & Charles Clifton. (1987). Interactive use of lexical information in speech perception. *Journal of Experimental Psychology: Human Perception and Performance* 13: 291-99.

Connine, Cynthia M., Debra Titone, Thomas Deelman & Dawn Blasko. (1997). Similarity mapping in spoken word recognition. *Journal of Memory and Language* 37: 463-80.

Crowley, Terry. (1998). The voiceless fricatives [s] and [h] in Erromangan: One phoneme, two, or one and a bit? *Australian Journal of Linguistics* 18: 149-68.

Cutler, Anne & Dennis Norris. (1979). Monitoring Sentence Comprehension. *Sentence Processing: Psycholinguistic Studies Presented to Merrill Garrett*, (eds) Cooper, W. E. & E. C. T. Walker, p.113-34. Hillsdale: Erlbaum.

Deelman, Thomas & Cynthia M. Connine. (2001). Missing information in spoken word recognition: Nonreleased stop consonants. *Journal of Experimental Psychology: Human Perception and Performance* 27: 656-63.

Dehaene-Lambertz, Ghislaine. (1997). Electrophysiological correlates of categorical phoneme perception in adults. *NeuroReport* 8: 919-24.

Dixon, Robert Malcolm Ward. (1970). Proto-Australian laminals. *Oceanic Linguistics* 9: 79-103.

Dong, Shaowen. (1958). *Yuyin Changtan [Introduction to Phonetics]*. Beijing: Wenhua Jiaoyu Chubanshe.

Duanmu, San. (2007). *The Phonology of Standard Chinese.* Oxford University Press.

Dunbar, Ewan & William J. Idsardi. (2010). Review on A Critical Introduction to Phonology. *Phonology* 27: 325-31.

Ernestus, Mirjam & Harald Baayen. (2007). Paradigmatic effects in auditory word recognition: The case of alternating voice in Dutch. *Language and Cognitive Processes* 22: 1-24.

Eulitz, Carsten & Aditi Lahiri. (2004). Neurobiological evidence for abstract phonological representations in the mental lexicon during speech recognition. *Journal of Cognitive Neuroscience* 16: 577-83.

Fox, Robert Allen. (1984). Effect of lexical status on phonetic categorization. *Journal of Experimental Psychology: Human Perception and Performance* 10: 526-40.

Ganong, William F. (1980). Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance* 6: 110-25.

Goldinger, Stephen D. (1996). Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 22: 1166-83.

Goldinger, Stephen D. (1997). Perception and production in an episodic lexicon. *Talker Variability in Speech Processing*, (eds) Johnson, Keith & John W. Mullennix, p.33-66. San Diego: Academic Press.

Goldinger, Stephen D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review* 105: 251-79.

Goldsmith, John. (1995). Phonological Theory. *The Handbook of Phonological Theory*, (ed.) Goldsmith, John, p.1-23. Cambridge, MA: Blackwell.

Gussenhoven, Carlos. (1999). Illustrations of the IPA: Dutch. *Handbook of the International Phonetic Association*, p.74-77. Cambridge: Cambridge University Press.

Hall, Kathleen Currie. (2009). *A Probabilistic Model of Phonological Relationships from Contrast to Allophony*. The Ohio State University.

Harnsberger, James D. (2001). The perception of Malayalam nasal consonants by Marathi, Punjabi, Tamil, Oriya, Bengali, and American English listeners: A multidimensional scaling analysis. *Journal of Phonetics* 29: 303-27.

Harris, Zellig S. (1951). *Methods in Structural Linguistics.* Chicago: University of Chicago Press.

Hartman, Lawton M. (1944). The segmental phonemes of the Peiping dialect. *Language* 20: 28-42.

Hay, Jennifer, Janet Pierrehumbert & Mary Beckman. (2004). Speech perception, well-formedness, and the statistics of the lexicon. *Papers in Laboratory Phonology VI* Cambridge: Cambridge University Press.

Hockett, Charles F. (1942). A system of descriptive phonology. *Language* 18: 3-21.

Hualde, José Ignacio. (2004). Quasi-phonemic contrasts in Spanish. *WCCFL 23: Proceedings of the 23rd West Coast Conference on Formal Linguistics*, (eds) Chand, Vineeta, Ann Kelleher, Angelo J. Rodriguez & Banjamin Schmeiser, p.374-98. Somerville, MA: Cascadilla Press.

Huang, Tsan. (2001). The interplay of perception and phonology in tone 3 sandhi in Chinese Putonghua. *Studies on the Interplay of Speech Perception and Phonology*, (eds) Hume, Elizabeth & Keith Johnson, p.23-42. Ohio State University.

Hume, Elizabeth & Keith Johnson. (2003). The impact of partial phonological contrast on speech perception. *15th International Congress of Phonetic Sciences*, p.2385-88.

Inkelas, Sharon. (1995). The consequences of optimization for underspecification. *Proceedings of the 25th Meeting of North East Linguistic Society* 25: 287-302.

Iverson, Gregory K. & Ahrong Lee. (2006). Perception of contrast in Korean loanword adaptation. *Korean Linguistics* 13: 49-87.

Jaeger, Jeri J. (1986). Concept formation as a tool for linguistic research. *Experimental Phonology*, (eds) Ohala, John J. & Jeri J. Jaeger, p.211-37. Orlando: Academic Press.

Johnson, Keith. (1997). Speech perception without speaker normalization. *Talker Variability in Speech Processing*, (eds) Johnson, Keith & John W. Mullennix, p.145-65. San Diego: Academic Press.

Johnson, Keith. (2005a). Decisions and mechanisms in exemplar-based phonology. *UC Berkeley Phonology Lab Annual Report 2005*: 289-311.

Johnson, Keith. (2005b). Decisions and mechanisms in exemplar-based phonology. *UC Berkeley Phonology Lab Annual Report*: 289-311.

Johnson, Keith. (2006). Resonance in an exemplar-based lexicon: The emergence of social identity and phonology. *Journal of Phonetics* 34: 485-99.

Johnson, Keith & Molly Babel. (2010). On the perceptual basis of distinctive features: Evidence from the perception of fricatives by Dutch and English speakers. *Journal of Phonetics* 38: 127-36.

Kager, René. (1999). *Optimality Theory.* Cambridge University Press.

Kazanina, Nina, Colin Phillips & William Idsardi. (2006). The influence of meaning on the perception of speech sounds. *Proceedings of the National Academy of Sciences of the United States of America* 103: 11381-86.

Kim, Hyunsoon. (2009). Korean adaptation of English affricates and fricatives in a feature-driven model of loanword adaptation. *Loan Phonology*, (eds) Calabrese, Andrea & W. Leo Wetzels, p.155-80. John Benjamins.

Krämer, Martin. (2004). Lexical economy, optimization and the indeterminacy of underlying representations. *Handout from ULCL Workshop 'The Lexicon in Optimality Theory'*, Leiden.

Kristoffersen, Gjert. (2000). *The Phonology of Norwegian.* Oxford: Oxford University.

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

Ladd, Robert D. (2006). "Distinctive phones" in surface representation. *Laboratory Phonology 8*, (eds) Goldstein, Louis M., D. H. Whalen & Catherin T. Best Berlin: Mouton de Gruyter.

Ladefoged, Peter & Ian Maddieson. (1996). *The Sounds of the World's Languages.* Cambridge: Blackwell.

Lasky, Robert E., Ann Syrdal-Lasky & Robert E. Klein. (1975). VOT discrimination by four to six and a half month old infants from Spanish environments. *Journal of Experimental Child Psychology* 20: 215-25.

Li, Fangfang. (2008). *The Phonetic Development of Voiceless Sibilant Fricatives in English, Japanese and Mandarin Chinese*. The Ohio State University.

Lin, Yenhwei. (1989). *Autosegmental Treatment of Segmental Processes in Chinese Phonology*. University of Texas, Austin.

Lisker, Leigh. (2001). Hearing the Polish sibilants [s š ś]: phonetic and auditory judgements. *Travaux du Cercle Linguistique de Copenhague XXXI. To honour Eli Fischer-Jørgensen*, (eds) Grønnum, Nina & Jørgen Rischel, p.226-38. C.A. Reitzel: Copenhagen.

Lisker, Leigh & Arthur S. Abramson. (1970). The voicing dimensions: Some experiments in comparative phonetics. Paper presented to the *Proceedings of the Sixth International Congress of Phonetic Sciences,* Prague Academia.

Luck, Steven J. (2005). *An Introduction to the Event-Related Potential Technique.* Cambridge: The MIT Press.

MacKain, Kristine S., Catherine T. Best & Winifred Strange. (1981). Categorical perception of English /r/ and /l/ by Japanese bilinguals. *Applied Psycholinguistics* 2: 369-90.

Massaro, Dominic & Michael Cohen. (1983). Phonological constraints in speech perception. *Perception & Psychophysics* 34: 338-48.

McCarthy, John J. & Alan Prince. (1995). *Faithfulness and Reduplicative Identity.* University of Massachusetts, Amherst.

McClelland, James L. & Jeffrey L. Elman. (1986). The TRACE model of speech perception. *Cognitive Psychology* 18: 1-86.

McGuire, Grant Leese. (2007). *Phonetic Category Learning*. Diss. The Ohio State University.

Miller, Joanne L. & Peter D. Eimas. (1995). Speech perception: from signal to word. *Annual Review of Psychology* 46: 467-92.

Moulton, Keir. (2003). Deep allophones in the Old English laryngeal system. *Toronto Working Papers in Linguistics* 20: 157-73.

Moulton, William G. (1962). *The Sounds of English and German, A Systematic Analysis of the Contrasts Between the Sound Systems.* Chicago: University of Chicago Press.

Näätänen, Risto. (2001). The perception of speech sound by the human brain as reflected by the mismatch negativity (MMN) and its magnetic equivalent (MMNm). *Psychophysiology* 38: 1-21.

Näätänen, Risto, Anne Lehtokoski, Mietta Lennes, Marie Cheour, Minna Huotilainen, Antti Livonen, Martti Valnio, Paavo Alku, Risto J. Ilmoniemi, Aavo Luuk, Jüri Allik, Janne Sinkkonen & Kimmo Alho. (1997). Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature* 385: 432-34.

Näätänen, Risto, Petri Paavilainen, T. Rinne & K. Alho. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clinical Neurophysiology* 118: 2544-90.

Norris, Dennis, James M. McQueen & Anne Culter. (2000). Merging information in speech recognition: Feedback is never necessary. *Behavioral And Brain Sciences* 23: 299-370.

Norris, Dennis, James M. McQueen & Anne Cutler. (2003). Perceptual learning in speech. *Cognitive Psychology* 47: 204-38.

Pierrehumbert, Janet B. (2001a). Exemplar dynamics: Word frequency, lenition, and contrast. *Frequency and the Emergence of Linguistics Structure*, (eds) Bybee, Joan L. & Paul Hopper, p.137-57. Philadelphia: John Benjamins.

Pierrehumbert, Janet B. (2001b). Stochastic phonology. *Glot International 5/6*, p.195-207. Oxford: Blackwell.

Pierrehumbert, Janet B. (2003a). Phonetic diversity, statistical learning, and acquisition of phonology. *Language and Speech* 46: 115-54.

Pierrehumbert, Janet B. (2003b). Probabilistic phonology: Discrimination and robustness. *Probabilistic Linguistics*, (eds) Bod, Rens, Jennifer Hay & Stefanie Jannedy, p.177-228. MIT Press.

Pierrehumbert, Janet B. (2006a). The next toolkit. *Journal of Phonetics* 34: 516-30.

Pierrehumbert, Janet B. (2006b). The statistical basis of an unnatural alternation. *Laboratory Phonology VIII, Varieties of Phonological Competence*, (eds) Goldstein, Louis, D. H. Whalen & Catherine T. Best, p.81-107. Berlin: Mouton de Gruyter.

Pisoni, David B. (1977). Identification and discrimination of the relative onset time of two component tones: Implications for voicing perception in stops. *Journal of Acoustic Society of America* 61: 1352-62.

Pisoni, David B. & Joan H. Lazarus. (1974). Categorical and noncategorical modes of speech perception along the voicing continuum. *The Journal of the Acoustical Society of America* 55: 328-33.

Pisoni, David B. & Jeffrey Tash. (1974). Reaction times to comparisons within and across phonetic categories. *Perception & Psychophysics* 15: 285-90.

Prince, Alan & Paul Smolensky. (1993). *Optimality Theory: Constraint Interaction in Generative Grammar*. Malden: Blackwell.

Pulleyblank, Edwin G. (1984). Vowelless Chinese? An application of the three-tiered theory of syllable structure. *Proceedings of the Sixteenth International Conference on Sino-Tibetan Languages and Linguistics*, p.586-619. Seattle: University of Washinton.

Ramsey, S. Robert. (1987). *The Languages of China.* Princeton: Princeton University Press.

Ranbom, Larissa J. & Cynthia M. Connine. (2007). Lexical representation of phonological variation in spoken word recognition. *Journal of Memory and Language* 57: 273-98.

Rose, Sharon & Lisa King. (2007). Speech error elicitation and co-occurrence restrictions in two Ethiopian semitic languages. *Language and Speech* 50: 451-504.

Scobbie, James M. & Jane Stuart-Smith. (2008). Quasi-phonemic contrast and the indeterminacy of the segmental inventory: Examples from Scottish English. *Contrast in Phonology: Perception and Acquisition*, (eds) Avery, Peter, B. Elan Dresher & Keren Rice Berlin: Mouton.

Shih, Chilin. (1997). Mandarin third tone sandhi and prosodic structure. *Studies in Chinese Phonology*, (ed.) Wang, Jialing, p.81-124. New York: Mouton de Gruyter.

Silverman, Daniel. (2006). *A Critical Introduction to Phonology of Sound, Mind, and Body.* London & New York: Continuum.

Sohn, Ho-Min. (1999). *The Korean Language.* Cambridge University Press.

Steinberg, Johanna, Hubert Truckenbrodt & Thomas Jacobsen. (2009). Preattentive phonotactic processing as indexed by the Mismatch Negativity. *Journal of Cognitive Neuroscience* 22: 2174-85.

Steinberg, Johanna, Hubert Truckenbrodt & Thomas Jacobsen. (2010). Activation and application of an obligatory phonotactic constraint in German during automatic speech processing is revealed by human event-related potentials. *International Journal of Psychophysiology* 56: 143-55.

Steriade, Donca. (2008). The phonology of perceptibility effect: The p-map and its consequences for constraint organization. *The Nature of the Word*, (eds) Hanson, Kristin & Sharon Inkelas, p.151-79. Cambridge: MIT Press.

Suh, Yunju. (2009a). Perception of English voiceless alveolar and postalveolar fricative before /i/ by Korean speakers. *Acoustical Society of America (ASA) 2nd ASA Special Workshop on Speech*, Portland.

Suh, Yunju. (2009b). *Phonological and Phonetic Asymmetries of Cw Combinations*. Stony Brook Stony Brook University.

Sumner, Meghan & Arthur G. Samuel. (2005). Perception and representation of regular variation: The case of final /t/. *Journal of Memory and Language* 52: 322-38.

Swadesh, Morris. (1934). The phonemic principle. *Language* 10: 117-29.

Trager, George L. (1934). The phonemes of Russian. *Language* 10: 334-44.

Trubetzkoy, Nikolai Sergeevich. (1969). *Principles of Phonology (Translated by Christiane A. M. Baltaxe).* Berkeley and Los Angeles: University of California Press.

Tung, Tung-ho. (1954). *Zhongguo Yuyin Shi [A Phonological History of Chinese]*. Taipei: Chonghua Wenhua Chubanshe.

Vennemann, Theo. (1971). The phonology of Gothic vowels. *Language* 47: 90-132.

Wan, I-Ping. (2010). Phonological experiments in the study of palatals in Mandarin. *Journal of Chinese Linguistics* 38: 157-74.

Warner, Natasha, Allard Jongman, Joan Sereno & Rachel Kemps. (2004). Incomplete neutralization and other sub-phonemic durational differences in production and perception: evidence from Dutch. *Journal of Phonetics* 32: 251-76.

Werker, Janet F. & Chris E. Lalonde. (1988). Cross-language speech perception: Initial capabilities and developmental change. *Developmental Psychology* 24: 672-83.

Wiese, Richard. (1997). Underspecification and the description of Chinese vowels. *Studies in Chinese Phonology*, (eds) Wang, Jialing & Norval Smith, p.219-49. Berlin: Mouton de Gruyter.

Wu, Yuwen. (1994). *Mandarin Segmental Phonology*. Toronto: University of Toronto.

Wurm, Lee H. & Arthur G. Samuel. (1997). Lexical inhibition and attentional allocation during speech perception: Evidence from phoneme monitoring. *Journal of Memory and Language* 36: 165-87.

Yip, Moira. (1996). Lexical optimization in languages without alternations. *Current Trends in Phonology: Models and methods*, (eds) Durand, J. & B. Laks, p.757-488. CNRS, Paris-X and University of Salford: University of Salford Publications.

Zsiga, Elizabeth C. (1995). An acoustic and electropalatographic study of lexical and post-lexical palatalization in American English. *Phonology and Phonetic Evidence: Papers in Laboratory Phonology IV*, (eds) Connell, B. & A. Arvaniti, p.282-302. Cambridge: Cambridge University Press.

Appendix A. English ability questionnaire

Participant number: ________________________

Email: ________________________________

Age: __________________________________

Gender: ______________________________

Language Group:

- What languages do you speak?

- Self-rated English ability

	Very bad						Very good
	1	2	3	4	5	6	7
Listening	☐	☐	☐	☐	☐	☐	☐
Speaking	☐	☐	☐	☐	☐	☐	☐
Reading	☐	☐	☐	☐	☐	☐	☐
Writing	☐	☐	☐	☐	☐	☐	☐
Overall	☐	☐	☐	☐	☐	☐	☐

Appendix B. English/Mandarin/Korean Semantic priming wordlist

B.1 English

	Prime	Target	Stimuli	Relation
1	Sunday	Monday	s	related
2	saddle	horse	s	related
3	samba	dance	s	related
4	sapful	tree	s	related
5	satan	devil	s	related
6	second	first	s	related
7	sender	mail	s	related
8	senior	junior	s	related
9	sentence	phrase	s	related
10	sequel	prequel	s	related
11	servant	master	s	related
12	sibling	brother	s	related
13	secret	whisper	s	related
14	sonic	hedgehog	s	related
15	suffer	pain	s	related
16	sultry	sexy	s	related
17	supply	demand	s	related
18	silver	gold	s	related
19	sabbath	bags	s	unrelated
20	sadden	holy	s	unrelated
21	salad	depress	s	unrelated
22	saucy	US	s	unrelated
23	safety	waltz	s	unrelated
24	segment	dressing	s	unrelated
25	session	beachball	s	unrelated
26	silent	evil	s	unrelated
27	sober	tomato	s	unrelated
28	subject	pin	s	unrelated
29	succeed	letter	s	unrelated

30	sullen	place	s	unrelated
31	summer	piece	s	unrelated
32	summon	ring	s	unrelated
33	symbol	verb	s	unrelated
34	sanction	smoke	s	unrelated
35	simple	running	s	unrelated
36	saying	boring	s	unrelated
37	shabby	shack	sh	related
38	shackle	chain	sh	related
39	shadow	darkness	sh	related
40	shepherd	sheep	sh	related
41	shinbone	leg	sh	related
42	shiver	cold	sh	related
43	shoulder	arm	sh	related
44	shouting	yelling	sh	related
45	shovel	dig	sh	related
46	shifty	eyes	sh	related
47	sherlock	holmes	sh	related
48	shop_front	window	sh	related
49	sharing	caring	sh	related
50	shimmy	shake	sh	related
51	shoddy	work	sh	related
52	shatter	glass	sh	related
53	shaker	salt	sh	related
54	sharpen	pencil	sh	related
55	shapeless	cowboy	sh	unrelated
56	shoelace	sodden	sh	unrelated
57	shotgun	cheese	sh	unrelated
58	shudder	police	sh	unrelated
59	shuffle	church	sh	unrelated
60	shading	pizza	sh	unrelated
61	shaking	crosswalk	sh	unrelated
62	shamble	measure	sh	unrelated
63	shanty	citizen	sh	unrelated
64	sheriff	book	sh	unrelated
65	sheep_dog	writing	sh	unrelated
66	sheepish	only	sh	unrelated
67	sugar	noise	sh	unrelated
68	shameless	warehouse	sh	unrelated
69	shelter	success	sh	unrelated
70	shutter	hot	sh	unrelated
71	shoofly	magic	sh	unrelated

72	shameful	relax	sh	unrelated
73	linking	notice	filler	word
74	medial	make	filler	word
75	looking	publish	filler	word
76	carry	world	filler	word
77	tunnel	nothing	filler	word
78	nasty	cute	filler	word
79	lecture	proven	filler	word
80	battle	pattern	filler	word
81	manage	boomer	filler	word
82	behind	road	filler	word
83	tattoo	gradient	filler	word
84	beauty	involve	filler	word
85	modern	future	filler	word
86	partial	filler	filler	word
87	body	acquire	filler	word
88	passage	Spain	filler	word
89	priceless	topic	filler	word
90	nothing	table	filler	word
91	purpose	corm	filler	pseudowords
92	massive	bartle	filler	pseudowords
93	pillar	pragle	filler	pseudowords
94	profile	daver	filler	pseudowords
95	cover	prend	filler	pseudowords
96	mortal	tuspy	filler	pseudowords
97	tonight	lapic	filler	pseudowords
98	table	zibble	filler	pseudowords
99	prepare	hame	filler	pseudowords
100	listen	jashly	filler	pseudowords
101	notice	fashy	filler	pseudowords
102	daybreak	dop	filler	pseudowords
103	morning	lork	filler	pseudowords
104	problem	timble	filler	pseudowords
105	custom	vink	filler	pseudowords
106	beneath	withyard	filler	pseudowords
107	building	gamper	filler	pseudowords
108	careful	reparn	filler	pseudowords
109	topic	unham	filler	pseudowords
110	ready	preed	filler	pseudowords
111	danger	billump	filler	pseudowords
112	conceal	thimmel	filler	pseudowords
113	torture	hobben	filler	pseudowords

114	posture	whickle	filler	pseudowords
115	belly	repind	filler	pseudowords
116	paper	jamed	filler	pseudowords
117	despite	mank	filler	pseudowords
118	narrow	plisher	filler	pseudowords
119	chuckled	tammock	filler	pseudowords
120	genius	detrite	filler	pseudowords
121	quarter	pogin	filler	pseudowords
122	begin	flottler	filler	pseudowords
123	toddler	ringuin	filler	pseudowords
124	complete	preel	filler	pseudowords
125	remove	drepo	filler	pseudowords
126	toward	fillger	filler	pseudowords
127	correct	linety	filler	pseudowords
128	patient	crinnet	filler	pseudowords
129	dozen	pastrel	filler	pseudowords
130	temples	hostire	filler	pseudowords
131	bottom	bluve	filler	pseudowords
132	regret	mixelle	filler	pseudowords
133	pattern	lorp	filler	pseudowords
134	meeting	gleep	filler	pseudowords
135	parent	vuggle	filler	pseudowords
136	really	pultace	filler	pseudowords
137	castle	wronk	filler	pseudowords
138	enter	boik	filler	pseudowords
139	presence	plinky	filler	pseudowords
140	painter	quimmed	filler	pseudowords
141	rather	linket	filler	pseudowords
142	tension	abmute	filler	pseudowords
143	private	oblimp	filler	pseudowords
144	maybe	infish	filler	pseudowords
145	perfect	seegre	filler	pseudowords
146	desert	ogril	filler	pseudowords
147	produce	unjim	filler	pseudowords
148	upright	outpill	filler	pseudowords
149	embrace	capdime	filler	pseudowords
150	utter	drim	filler	pseudowords
151	appear	clagger	filler	pseudowords
152	concept	maggle	filler	pseudowords
153	request	unkim	filler	pseudowords
154	escape	aspill	filler	pseudowords
155	bathroom	zeen	filler	pseudowords

156	barely	brupy	filler	pseudowords
157	affect	ropple	filler	pseudowords
158	output	krinky	filler	pseudowords
159	order	obdupe	filler	pseudowords
160	transform	pitkin	filler	pseudowords
161	observe	linreed	filler	pseudowords
162	wooden	treeple	filler	pseudowords
163	prayer	skince	filler	pseudowords
164	ancient	sadpin	filler	pseudowords
165	author	jupin	filler	pseudowords
166	able	kilpat	filler	pseudowords
167	even	grettale	filler	pseudowords
168	crystal	brigga	filler	pseudowords
169	concern	prand	filler	pseudowords
170	writing	gannet	filler	pseudowords
171	translate	whabed	filler	pseudowords
172	unstable	vindle	filler	pseudowords
173	become	hollut	filler	pseudowords
174	cancer	hing	filler	pseudowords
175	only	dretty	filler	pseudowords
176	partner	gonale	filler	pseudowords
177	open	potbill	filler	pseudowords
178	promise	retupe	filler	pseudowords
179	open	linreed	filler	pseudowords
180	promise	treeple	filler	pseudowords

B.2 Mandarin

	Prime		Target		Stimuli	Relation
1	習慣	xiguan	自然	ziran	sh	related
2	西瓜	xigua	夏天	xiatian	sh	related
3	稀有	xiyou	珍貴	zhengue	sh	related
4	嘻笑	xixiao	怒罵	numa	sh	related
5	吸管	xiguan1	飲料	yinliao	sh	related
6	息怒	xinu	生氣	shengqi	sh	related
7	西方	xifang	東方	dongfang	sh	related
8	媳婦	xifu	婆婆	puopuo	sh	related
9	喜宴	xiyan	結婚	jiehun	sh	related
10	興趣	xinchu	喜歡	xihuan	sh	related
11	薪水	xinshui	錢	qian	sh	related

12	星光	xinguang	大道	dadao	sh	related
13	信仰	xinyang	宗教	zonjiao	sh	related
14	夕陽	xiyang	西下	xixia	sh	related
15	熄火	xihuo	拋錨	paomao	sh	related
16	心情	xinqing	愉快	yukuai	sh	related
17	欣賞	xinshang	美女	meinu	sh	related
18	行為	xingwei	舉止	juzhi	sh	related
19	希望	xiwang	以前	yichien	sh	unrelated
20	吸引	xiyin	石頭	shetou	sh	unrelated
21	洗澡	xizao	記者	jizhe	sh	unrelated
22	溪水	xishuei	工作	gongzuo	sh	unrelated
23	昔日	xire	血型	xuexing	sh	unrelated
24	隙縫	xifeng	朋友	pengyou	sh	unrelated
25	新聞	xinwen	星星	xinxin	sh	unrelated
26	辛苦	xinku	制式	zhishi	sh	unrelated
27	信任	xinren	罪犯	zuefan	sh	unrelated
28	星座	xinzuo	端莊	duanzhuang	sh	unrelated
29	形象	xinxiang	電話	dianhua	sh	unrelated
30	型式	xinshe	白色	baise	sh	unrelated
31	刑責	xinze	自信	zixin	sh	unrelated
32	幸運	xingyun	後悔	hauhue	sh	unrelated
33	姓名	xingmin	卜派	pupai	sh	unrelated
34	杏仁	xinren1	指教	zhejiao	sh	unrelated
35	信心	xinxing	替身	tishen	sh	unrelated
36	醒悟	xingwu	政治	zhenzhe	sh	unrelated
37	四維	sewe	八德	bade	s	related
38	似乎	sehu	好像	haoxiang	s	related
39	飼料	seliao	雞	ji	s	related
40	私立	seli	學校	xuexiau	s	related
41	撕票	sepiao	綁架	bangjia	s	related
42	思量	seliang	考慮	kaulu	s	related
43	祀奉	sefong	神明	shenmin	s	related
44	四川	sechuan	麻辣	mala	s	related
45	撕破	sepuo	臉	lian	s	related
46	飼養	seyang	寵物	chongwu	s	related
47	寺廟	semiao	和尚	heshang	s	related
48	斯文	sewen	眼鏡	yangjing	s	related
49	絲綢	sechou	絲路	selu	s	related
50	私有	seyou	財產	caichan	s	related
51	死刑	sesing	犯人	fanren	s	related

52	肆虐	senue	颱風	taifeng	s	related
53	四肢	sezhi	發達	fada	s	related
54	四面	semian	楚歌	chuge	s	related
55	思考	sekau	神明	shenming	s	unrelated
56	死亡	sewang	秘密	mimi	s	unrelated
57	司法	sefa	手帕	shoupa	s	unrelated
58	司機	seji	菜瓜	caigua	s	unrelated
59	思念	senian	軍人	junren	s	unrelated
60	賜給	segei	盲點	mangdian	s	unrelated
61	四處	sechu	條約	tiaoyue	s	unrelated
62	賜福	sefu	公開	gongkai	s	unrelated
63	私下	sexia	老師	laoshe	s	unrelated
64	絲巾	sejin	爭戰	zhanzheng	s	unrelated
65	絲瓜	segua	分手	fenshou	s	unrelated
66	司令	seling	大腦	danau	s	unrelated
67	死角	sejiao	屍體	sheti	s	unrelated
68	私人	seren	到處	dauchu	s	unrelated
69	撕毀	sehue	駕駛	jiashe	s	unrelated
70	死心	sesin	家人	jiaren	s	unrelated
71	賜教	sejiao1	禮物	liwu	s	unrelated
72	廝殺	sesha	機關	jiguang	s	unrelated
73	菠菜	buocai	媽媽	mama	filler	word
74	批評	piping	不入	buru	filler	word
75	替代	tidai	雪景	xuejin	filler	word
76	鬥爭	douzheng	考試	kaushe	filler	word
77	嘮叨	laodao	聯考	liankau	filler	word
78	高中	gaozhong	不同	butong	filler	word
79	刀槍	daoqiang	沖沖	chongchong	filler	word
80	筆記	biji	地區	diqu	filler	word
81	路燈	ludeng	嘉賓	jiabin	filler	word
82	地區	diqu	進取	jinchu	filler	word
83	怒氣	nuqi	觀眾	guanzhong	filler	word
84	客人	keren	頭髮	toufa	filler	word
85	特別	tebie	互助	huzhu	filler	word
86	樂觀	leguang	夢想	mengxiang	filler	word
87	各位	gewei	異性	yixing	filler	word
88	貓王	maowang	睡覺	shuejiao	filler	word
89	幫忙	bangmang	烤肉	kaurou	filler	word
90	美麗	meili	大兵	dabing	filler	word
91	逼迫	bipuo	比張	bizhang	filler	pseudoword

92	靠近	kaojin	爹身	dieshen	filler	pseudoword
93	保管	baoguang	停行	tingxing	filler	pseudoword
94	當場	dangchang	卡上	kashang	filler	pseudoword
95	賠償	peichang	金棉	jingmian	filler	pseudoword
96	腦筋	naojin	通瓶	tongpin	filler	pseudoword
97	固定	guding	螺拼	luopin	filler	pseudoword
98	旁觀	pangguan	究斗	chioudou	filler	pseudoword
99	偷竊	touqie	漂偷	piaotou	filler	pseudoword
100	密集	miji	淵傳	yuanchuan	filler	pseudoword
101	革命	geming	國舖	guopu	filler	pseudoword
102	泥土	nitu	丟引	diouyin	filler	pseudoword
103	禮貌	limao	郭到	guodao	filler	pseudoword
104	能力	nengli	白能	bainen	filler	pseudoword
105	棉被	mianbei	審頭	shendou	filler	pseudoword
106	拋棄	paoqi	姑次	guci	filler	pseudoword
107	湯匙	tangche	包搭	baoda	filler	pseudoword
108	泡沫	paomuo	跟計	genji	filler	pseudoword
109	包裝	baozhuang	鑾跑	luanpao	filler	pseudoword
110	米酒	mijou	央下	yangxia	filler	pseudoword
111	德國	deguo	禿角	tujiao	filler	pseudoword
112	湯圓	tangyuan	紅播	hongbuo	filler	pseudoword
113	奴隸	nuli	苦在	kucai	filler	pseudoword
114	勞動	laodong	工盆	gongpen	filler	pseudoword
115	北部	beibu	殘不	canbu	filler	pseudoword
116	痞子	pizi	英空	yingkong	filler	pseudoword
117	某年	mounian	某湃	moubai	filler	pseudoword
118	斗六	douliou	談狗	tangou	filler	pseudoword
119	推動	tuedong	腎色	shense	filler	pseudoword
120	地板	dipan	坑節	kengjie	filler	pseudoword
121	拉扯	lache	逼服	bifu	filler	pseudoword
122	剛強	kangqiang	龍改	longgai	filler	pseudoword
123	康慨	kangkai	判番	panfan	filler	pseudoword
124	兵器	binqi	卡嘆	katan	filler	pseudoword
125	拼命	pinming	詹郎	zhanlang	filler	pseudoword
126	猛烈	menglie	空鼻	kongbi	filler	pseudoword
127	董事	dongshi	看力	kanli	filler	pseudoword
128	聽取	tingqu	個探	getan	filler	pseudoword
129	隆重	longzhong	溫離	wenli	filler	pseudoword
130	功課	gongke	乾中	ganzhong	filler	pseudoword
131	空氣	kongqi	哭西	kuxi	filler	pseudoword

132	標示	biaoshi	魯標	lubiao	filler	pseudoword
133	瀑布	pubu	蕩關	dangguan	filler	pseudoword
134	幕僚	muliao	政印	zhengyin	filler	pseudoword
135	賭場	duchang	躺盤	tangpan	filler	pseudoword
136	圖片	tupian	砍色	kanse	filler	pseudoword
137	烈酒	liejou	高旁	gaopang	filler	pseudoword
138	苟且	gouqie	歐深	oshen	filler	pseudoword
139	抗議	kangyi	赤威	chiwei	filler	pseudoword
140	比較	bijiao	不胼	bupian	filler	pseudoword
141	判斷	panduan	是類	shilei	filler	pseudoword
142	漫長	manchang	扛業	kangyie	filler	pseudoword
143	冬天	dongtian	臘約	layue	filler	pseudoword
144	童年	tongnian	往畢	wangbi	filler	pseudoword
145	表態	biaotai	下股	xiagu	filler	pseudoword
146	統一	tongyi	頂安	dingan	filler	pseudoword
147	看法	kangfa	方傳	fangchuan	filler	pseudoword
148	更換	genghuan	被以	beiyi	filler	pseudoword
149	民眾	minchong	險罪	xiancue	filler	pseudoword
150	公道	kongdao	告看	gaokan	filler	pseudoword
151	騙人	pianren	水選	shuexuan	filler	pseudoword
152	目前	muqian	是許	shexu	filler	pseudoword
153	壟斷	longduan	且求	chiechiou	filler	pseudoword
154	動態	dongtai	棒首	bangshou	filler	pseudoword
155	通過	tongguo	巷放	xiangfang	filler	pseudoword
156	模式	muoshe	天只	tianzhe	filler	pseudoword
157	虧損	kuesun	在成	caicheng	filler	pseudoword
158	動物	dongwu	場比	changbi	filler	pseudoword
159	褓母	baomu	機到	jidao	filler	pseudoword
160	良好	lianghow	過社	guoshe	filler	pseudoword
161	邊境	bianchin	所影	suoyin	filler	pseudoword
162	匿名	nimin	被雄	beixiong	filler	pseudoword
163	公民	gongmin	打家	dajia	filler	pseudoword
164	悲傷	beshang	局用	juyong	filler	pseudoword
165	功效	kongxiao	拉米	lami	filler	pseudoword
166	台灣	taiwan	間搭	jianda	filler	pseudoword
167	待遇	daiyu	為地	wedi	filler	pseudoword
168	片面	pianmian	力不	libu	filler	pseudoword
169	肯定	kending	範且	fanqie	filler	pseudoword
170	大餅	dabing1	市選	shexuan	filler	pseudoword
171	排行	paihang	受這	shouzhe	filler	pseudoword

172	蘋果	pinguo	散國	sanguo	filler	pseudoword
173	提前	tiqian	萬否	wanfou	filler	pseudoword
174	敲定	chiaoding	序抗	xukang	filler	pseudoword
175	關鍵	guanjian	轍倒	chedao	filler	pseudoword
176	討論	taolun	聞度	wendu	filler	pseudoword
177	議題	yiti	擬休	nixiou	filler	pseudoword
178	嚴重	yanchong	推方	tuefang	filler	pseudoword
179	凌晨	lingchen	容活	ronghuo	filler	pseudoword
180	公共	gonggong	再客	caike	filler	pseudoword

B.3 Korean

	Prime	Target	Stimuli	Relation
1	승리	깊이	s	unrelated
2	승진	단풍	s	unrelated
3	승부	토마토	s	unrelated
4	습기	배우	s	unrelated
5	승계	크기	s	unrelated
6	스물	극본	s	unrelated
7	승화	팔	s	unrelated
8	습득	바람	s	unrelated
9	습격	편지	s	unrelated
10	승차	피부	s	unrelated
11	스릴	가게	s	unrelated
12	승격	겨울	s	unrelated
13	습지	꼬리	s	unrelated
14	승복	대문	s	unrelated
15	승률	두부	s	unrelated
16	스낵	마음	s	unrelated
17	스윙	반지	s	unrelated
18	습작	빨래	s	unrelated
19	시각	개미	sh	unrelated
20	시내	계란	sh	unrelated
21	시민	그네	sh	unrelated
22	시선	낮잠	sh	unrelated
23	시설	단어	sh	unrelated
24	시월	렌즈	sh	unrelated
25	시위	만남	sh	unrelated

26	시인	콩	sh	unrelated
27	시장	탈출	sh	unrelated
28	식사	파도	sh	unrelated
29	식탁	구두	sh	unrelated
30	식품	길이	sh	unrelated
31	신경	노래	sh	unrelated
32	신문	눈물	sh	unrelated
33	신분	도로	sh	unrelated
34	신청	바다	sh	unrelated
35	신호	마당	sh	unrelated
36	심장	트럭	sh	unrelated
37	스타	별	s	related
38	습관	버릇	s	related
39	스승	제자	s	related
40	슬픔	기쁨	s	related
41	스님	절	s	related
42	승객	버스	s	related
43	슬쩍	소매치기	s	related
44	승소	재판	s	related
45	슬기	지혜	s	related
46	스키	눈	s	related
47	승선	배	s	related
48	승자	패자	s	related
49	승낙	허락	s	related
50	습도	온도	s	related
51	스푼	포크	s	related
52	슬하	자식	s	related
53	승천	용	s	related
54	승마	말	s	related
55	시계	시간	sh	related
56	시골	농촌	sh	related
57	시작	끝	sh	related
58	식구	가족	sh	related
59	식당	밥	sh	related
60	식량	쌀	sh	related
61	식물	꽃	sh	related
62	시험	성적	sh	related
63	신고	경찰	sh	related
64	신발	구두	sh	related
65	신부	신랑	sh	related

66	신앙	종교	sh	related
67	신용	카드	sh	related
68	신체	건강	sh	related
69	신화	전설	sh	related
70	실내	실외	sh	related
71	실수	잘못	sh	related
72	실험	과학자	sh	related
73	경제	대야	filler	word
74	과일	공학	filler	word
75	기계	양배추	filler	word
76	껍질	나라	filler	word
77	남자	추수	filler	word
78	농민	감성	filler	word
79	느낌	입시	filler	word
80	도전	종점	filler	word
81	도착	친구	filler	word
82	로봇	안주	filler	word
83	맥주	꼬리	filler	word
84	모자	달성	filler	word
85	목표	총	filler	word
86	바보	대륙	filler	word
87	부엌	줄기	filler	word
88	뿌리	크림	filler	word
89	토론	오징어	filler	word
90	평화	야구	filler	word
91	감독	메걸	filler	pseudoword
92	고무	누패	filler	pseudoword
93	공부	두랑	filler	pseudoword
94	김치	기굽	filler	pseudoword
95	꽃밭	나소구	filler	pseudoword
96	나무	방두리	filler	pseudoword
97	날개	양준	filler	pseudoword
98	대학	제밤	filler	pseudoword
99	동네	찬태	filler	pseudoword
100	동물	타삭	filler	pseudoword
101	동생	표든	filler	pseudoword
102	머리	계민	filler	pseudoword
103	무기	놀보루	filler	pseudoword
104	발견	담채미	filler	pseudoword
105	병원	막반	filler	pseudoword

106	부모	도날	filler	pseudoword
107	커피	민반	filler	pseudoword
108	태풍	온사	filler	pseudoword
109	위기	제극	filler	pseudoword
110	초록	태북	filler	pseudoword
111	여자	출숭	filler	pseudoword
112	정신	포랑	filler	pseudoword
113	농장	각잔	filler	pseudoword
114	조립	농당	filler	pseudoword
115	머리	달민	filler	pseudoword
116	천재	막생	filler	pseudoword
117	의사	폑	filler	pseudoword
118	자식	굼	filler	pseudoword
119	엄마	눅	filler	pseudoword
120	여름	악잔	filler	pseudoword
121	티비	제고포	filler	pseudoword
122	가지	참댁	filler	pseudoword
123	인형	파말	filler	pseudoword
124	기타	더두리	filler	pseudoword
125	나물	가솝	filler	pseudoword
126	팔찌	내몽	filler	pseudoword
127	복도	대날	filler	pseudoword
128	머슴	박체	filler	pseudoword
129	난리	촘	filler	pseudoword
130	지도	야온	filler	pseudoword
131	안과	멸빛	filler	pseudoword
132	책상	라디건	filler	pseudoword
133	파랑	지액	filler	pseudoword
134	케찹	펄대	filler	pseudoword
135	땀샘	촉사	filler	pseudoword
136	거수	느막	filler	pseudoword
137	내기	국지장	filler	pseudoword
138	덧셈	일론	filler	pseudoword
139	리본	디담	filler	pseudoword
140	민물	초린	filler	pseudoword
141	모습	날성	filler	pseudoword
142	불안	봉각	filler	pseudoword
143	유림	책당	filler	pseudoword
144	재미	각솜	filler	pseudoword
145	충동	배걸	filler	pseudoword

146	피해	널만	filler	pseudoword
147	투자	잔티	filler	pseudoword
148	구강	자슴	filler	pseudoword
149	누리	농나문	filler	pseudoword
150	두부	걸강	filler	pseudoword
151	루머	뎁사	filler	pseudoword
152	매진	여목	filler	pseudoword
153	방송	밤채	filler	pseudoword
154	안경	병화기	filler	pseudoword
155	준비	칼륨	filler	pseudoword
156	축대	은박자	filler	pseudoword
157	칼집	걱황	filler	pseudoword
158	타격	민쇄	filler	pseudoword
159	폐해	몀	filler	pseudoword
160	규칙	둑생	filler	pseudoword
161	뉴스	참식	filler	pseudoword
162	독성	망화	filler	pseudoword
163	린스	닥배	filler	pseudoword
164	마진	포랑	filler	pseudoword
165	봉쇄	티비오	filler	pseudoword
166	양산	차온	filler	pseudoword
167	저울	올잎	filler	pseudoword
168	차이	볼잡	filler	pseudoword
169	킬로	낙파	filler	pseudoword
170	토끼	옴상	filler	pseudoword
171	포도	탕문	filler	pseudoword
172	교정	카테나	filler	pseudoword
173	노점	겸진	filler	pseudoword
174	다리	로방	filler	pseudoword
175	멧돌	장규	filler	pseudoword
176	보전	냉졸	filler	pseudoword
177	아들	곰기	filler	pseudoword
178	주식	주프	filler	pseudoword
179	처가	와단	filler	pseudoword
180	장갑	맥본	filler	pseudoword

Appendix C. English semantic priming wordlist (real words shaded)

English			
sh word		*s* word	
shabby	shapeless	Sunday→fun day	sabbath
shackle	shoelace	saddle	sadden
shadow	shotgun	samba	salad
shepherd	shudder	sapful	saucy
shinbone	shuffle	satan	safety
shiver	Shading→fading	second	segment
Shoulder→folder	Shaking→faking	Sender→fender	session
shouting	shamble	senior	silent
shovel	shanty	sentence	sober
shifty	sheriff	sequel	subject
sherlock	sheep dog	servant	succeed
shop front	sheepish	sibling	sullen
sharing	sugar	secret	summer
shimmy	Shameless→fameless	Sonic→phonic	summon
shoddy	Shelter→filter	suffer	symbol
Shatter→fatter	shutter	sultry	sanction
shaker	shaker	supply	simple
sharpen	shameful	silver	saying

Appendix D. Log transformed results on semantic priming

D.1 Mandarin

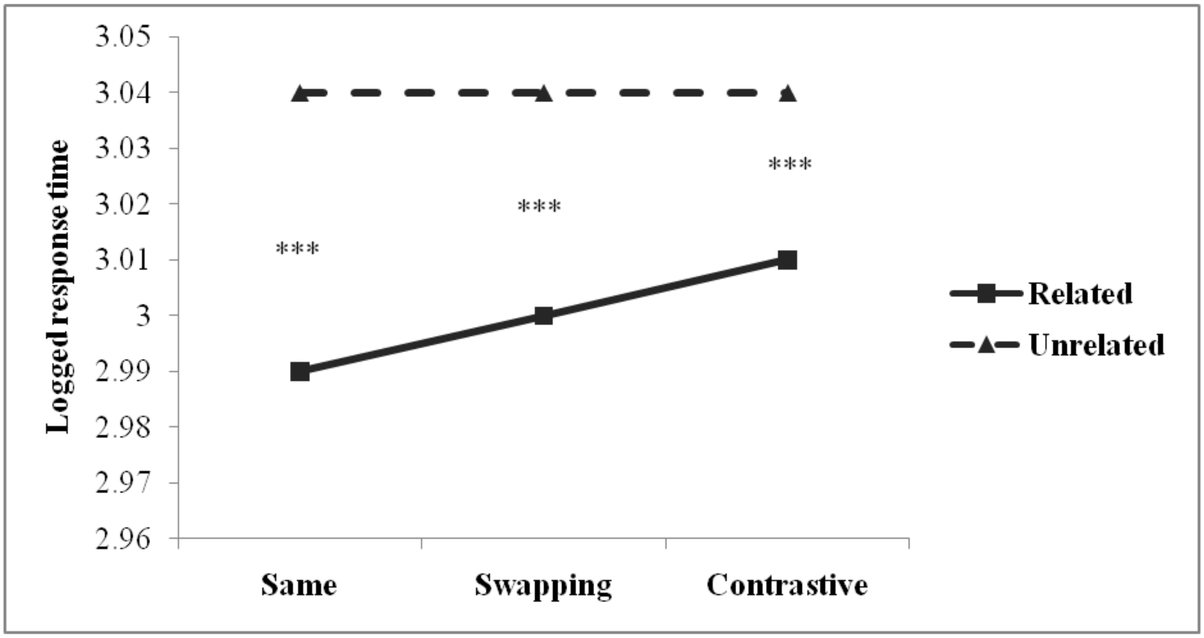
The mean logged RTs and the priming effects (the difference between Related and Unrelated) for the three experimental conditions are shown in Table D-1 with standard deviations in parentheses, and are illustrated in Figure D-1. (The *x* axis represents the different conditions and the y axis represents the logged response times; *: p<.05; **: p<.01; ***: p<.001; *n.s.*: not significant).

Table D-1 Mandarin lexical decision logged response time

Condition / Relation	Same	Swapping	Contrastive
related	2.99 (.045)	3.00 (.04)	3.01 (.05)
Unrelated	3.04 (.04)	3.04 (.04)	3.04 (.05)
Priming effect	.05	.04	.03

A two-way ANOVA was run (**Condition**: Same, Swapping, Contrastive **x Relation**: related or Unrelated) across participants. There was a main effect of Relation (F(1, 57) = 187.498, p<.001). Planned comparisons showed that targets preceded by related primes were identified more quickly than unrelated primes in all three conditions (Same F(1, 19)=76.573, p<.001; Swapping F(1, 19)=51.882, p<.001; Contrastive F(1, 19)=77.189, p<.001). Simple effect of Condition in related was not significant ($F(2, 60)=1.125$, $p=.332$) nor does the simple effect of Condition in Unrelated ($F(2, 60)=.034$, $p=.967$).

Figure D-1   Mandarin lexical decision logged response time



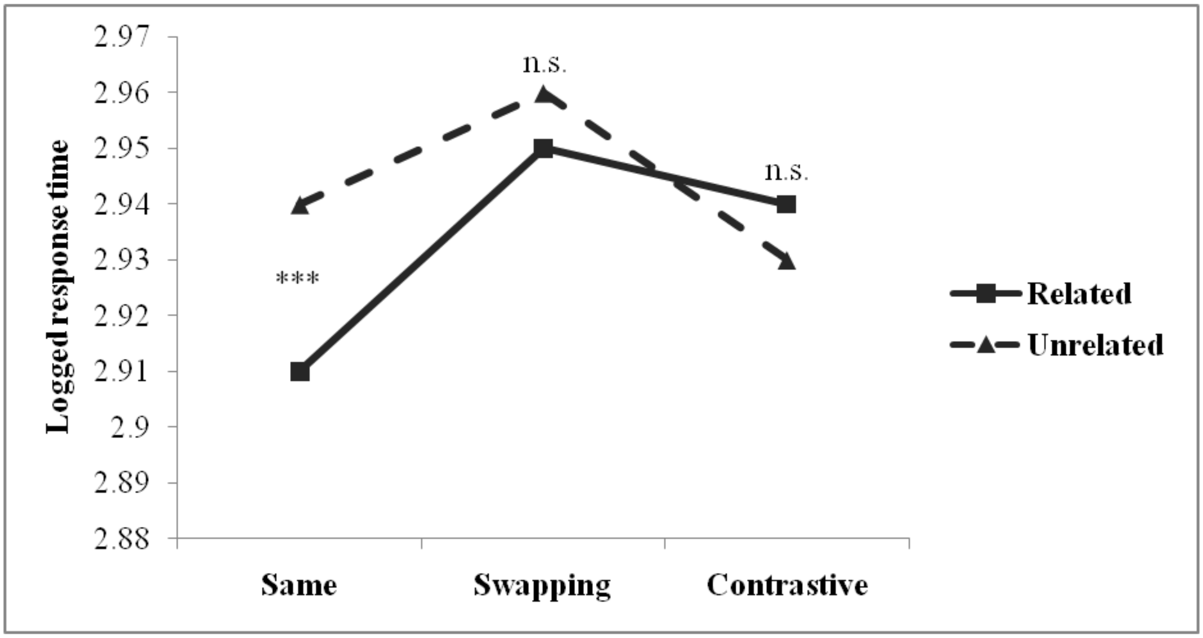
D.2 English

The mean logged RTs and the priming effects (the difference between Related and Unrelated) for the three experimental conditions are shown in Table D-2 with standard deviations in parentheses, and are illustrated in Figure D-2.

Table D-2   English lexical decision logged response time

Relation \ Condition	Same	Swapping	Contrastive
related	2.91 (.05)	2.95 (.05)	2.94 (.04)
Unrelated	2.94 (.05)	2.96 (.05)	2.93 (.04)
Priming effect	.03	.01	-.01

Figure D-2 English lexical decision logged response time



A two-way ANOVA (**Condition**: Same, Swapping, Contrastive **x Relation**: related or Unrelated) was run across participants. There is also an significant interaction ($F(2, 57)=8.824$, $p<.001$). Planned comparisons showed that only in the Same condition, targets preceded by the related primes were identified more quickly than the unrelated primes ($F(1,19)=73.707$, $p<.001$), but not in the other two conditions (Swapping $F(1,19)=2.657$, $p=.120$; Contrastive $F(1,19)=.560$, $p=.464$).

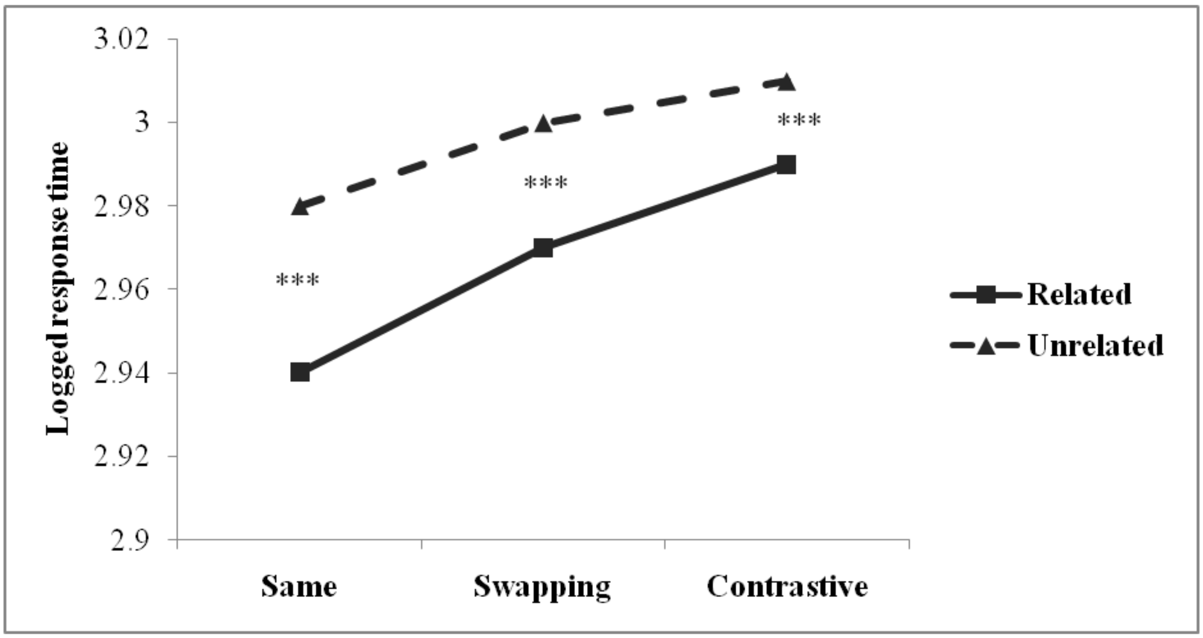
D.3 Korean

The mean logged RTs and the priming effects (the difference between Related and Unrelated) for the three experimental conditions are shown in Table D-3 with standard deviations in parentheses, and are illustrated in Figure D-3.

Table D-3 Korean logged lexical decision response time

Relation \ Condition	Same	Swapping	Contrastive
related	2.94 (.038)	2.97 (.059)	2.99 (.048)
Unrelated	2.98 (.036)	3.00 (.066)	3.01 (.048)
Priming effect	.04	.03	-.01

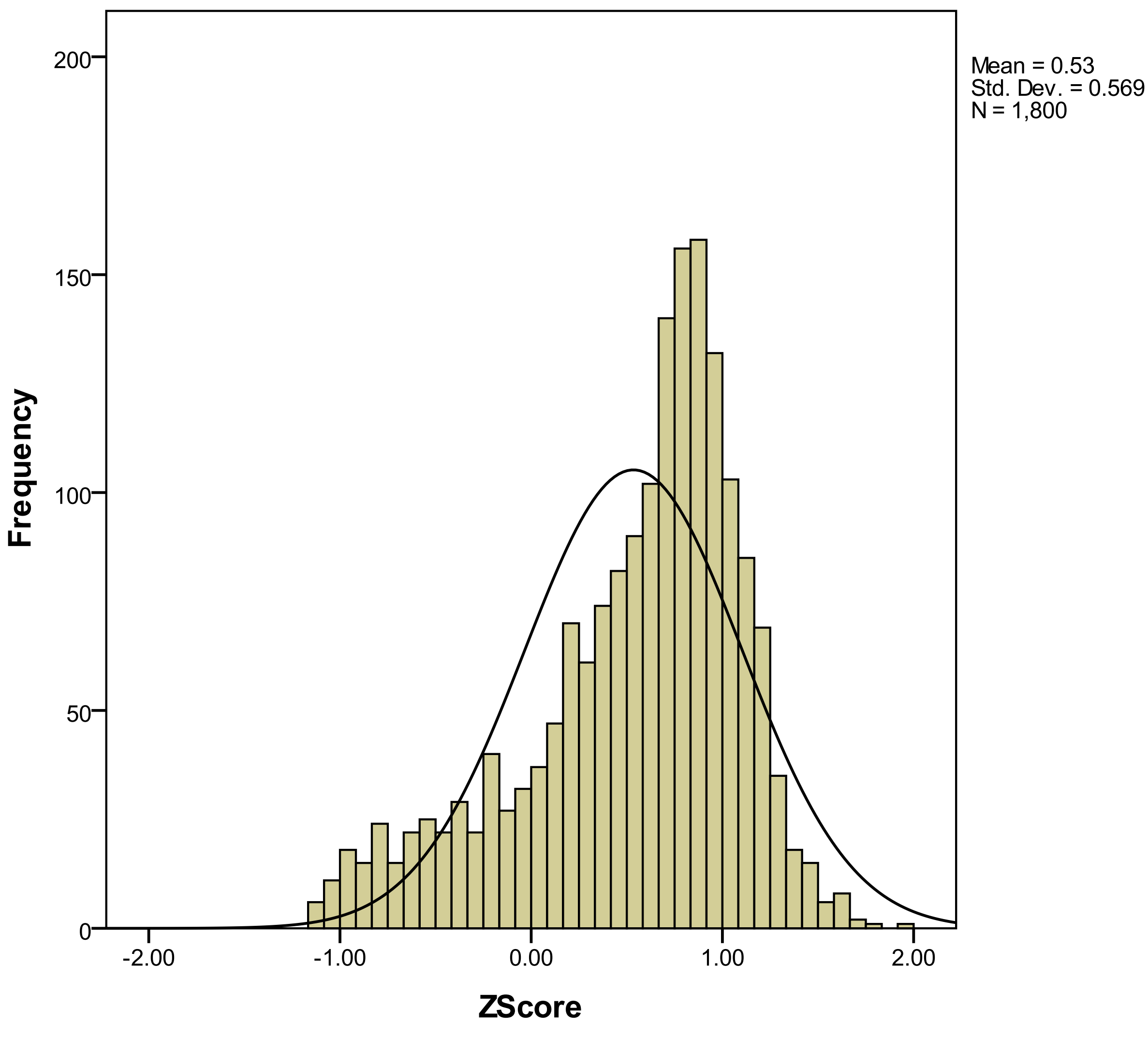
Figure D-3 Korean logged lexical decision response time



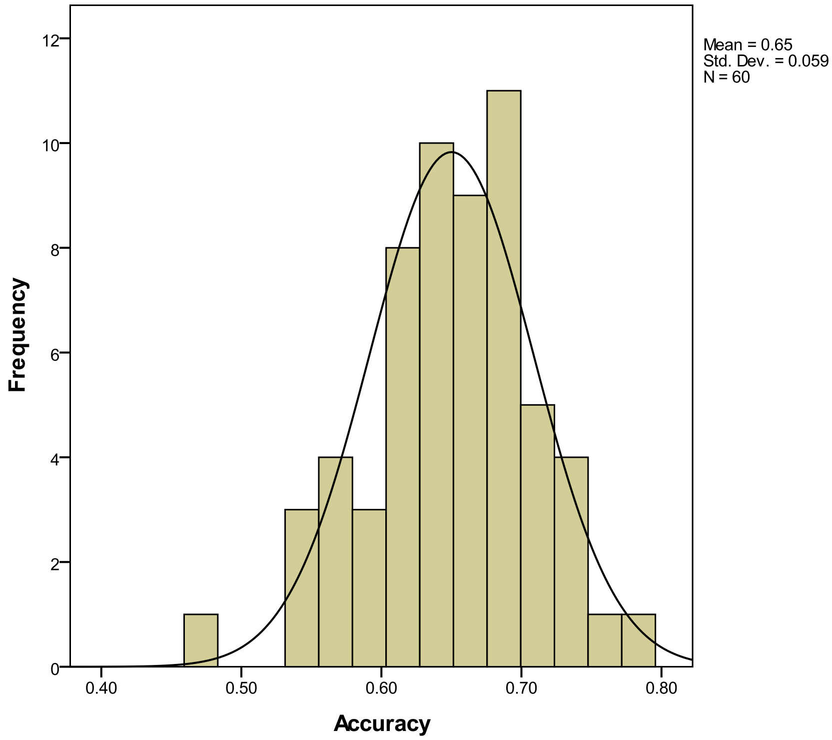
A two-way ANOVA was run (**Condition**: Same, Swapping, Contrastive **x Relation**: related or Unrelated) across participants. There was a main effect of Relation ($F(1, 57) = 128.668$, $p<.001$), and of Condition ($F(1, 57)=3.327$, $p<.05$. Planned comparisons showed that targets preceded by the related primes were identified more quickly than the unrelated primes in all three conditions (Same $F(1, 19)=78.801$, $p<.001$; Swapping $F(1, 19)=44.799$, $p<.001$; Contrastive $F(1, 19)=19.542$, $p<.001$). Simple effect of Condition in related was significant ($F(2, 60)=4.581$, $p<.05$). The significance is driven by the difference between the Same vs. Contrastive conditions (p=.013). The simple effect of Condition in Unrelated was not significant ($F(2, 60)=1.684$, $p=.195$).

Appendix E. Distributions of dependent variables

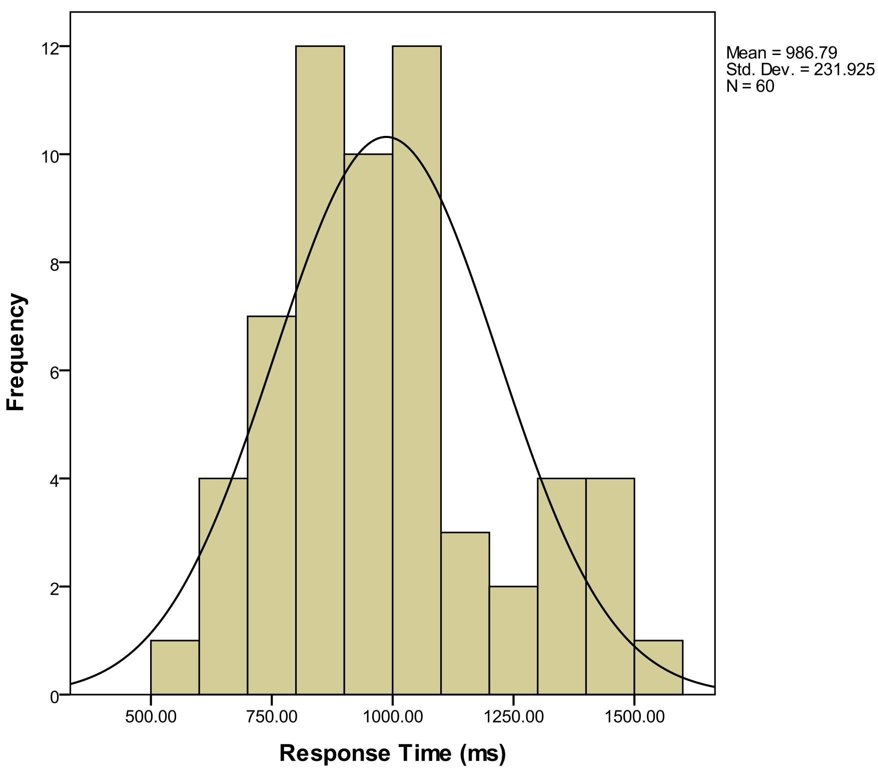
E.1 Distribution of the similarity rating z-scores



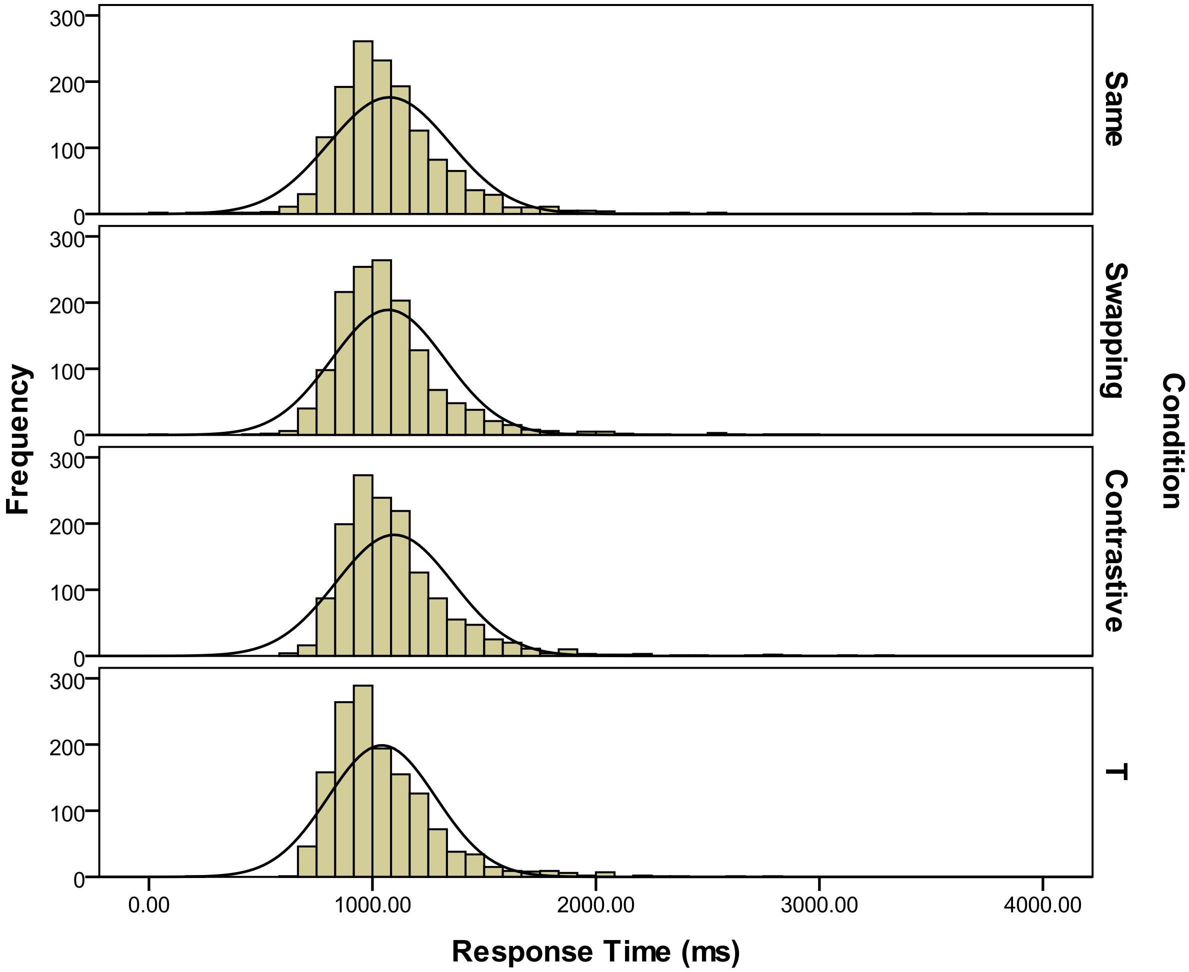
E.2 Distribution of the discrimination accuracy



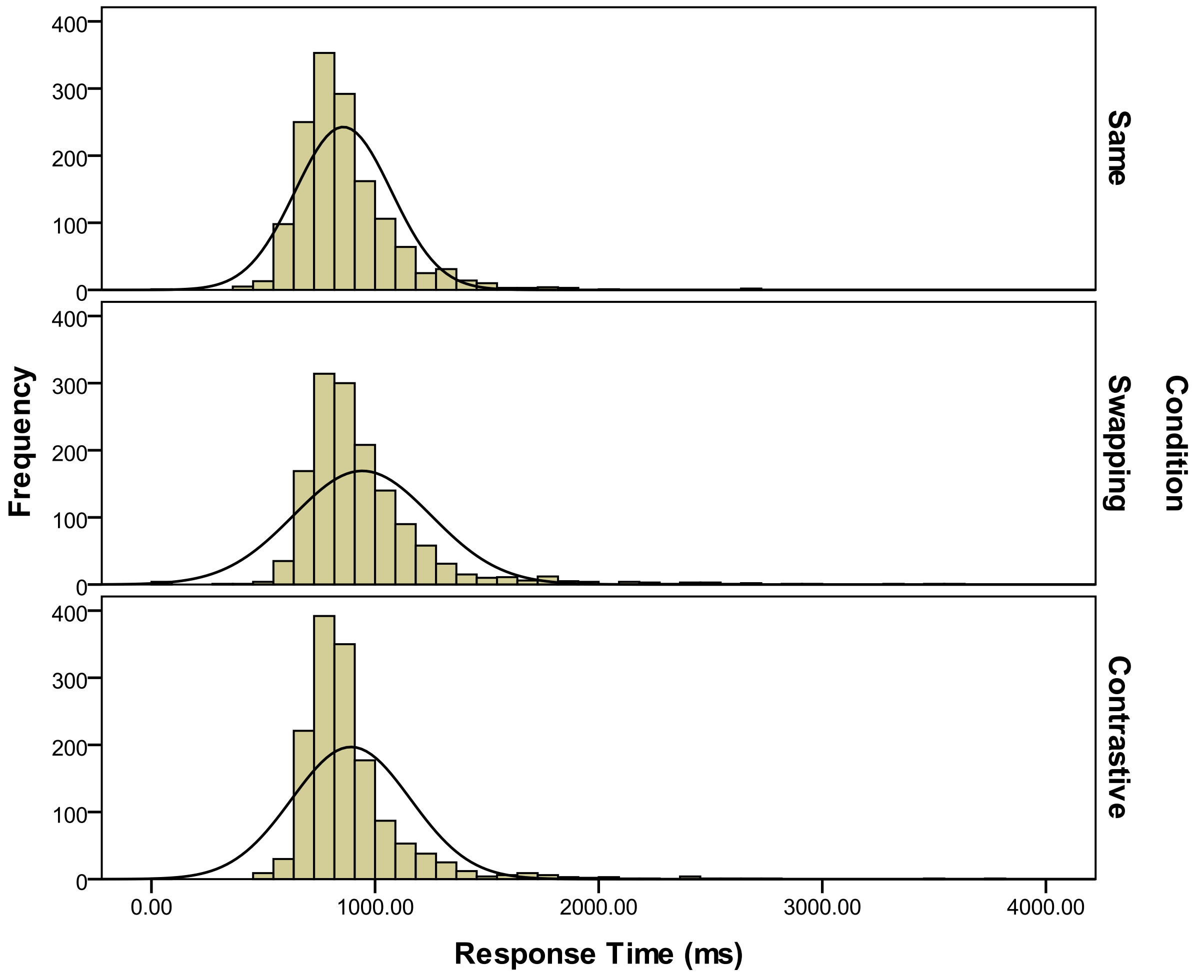
E.3 Distribution of the discrimination response time



E.4 Distribution of the Mandarin semantic priming lexical decision



E.5 Distribution of the English semantic priming lexical decision



E.6 Distribution of the Korean semantic priming lexical decision