

Testing the Roles of Distribution and Alternation in Phonological Relationships*

Yu-An Lu

National Chiao Tung University

Language and Linguistics
17(4) 501–528
© The Author(s) 2016
Reprints and permissions:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/1606822X16632113
lin.sagepub.com



This study investigates how two factors, distribution and morphological alternation, affect speakers' ability to group sounds into the same phoneme category. Previous findings indicate that allophonic variants of a single phoneme are rated as more similar than sounds belonging to separate phonemes. The present study builds on these findings by conducting a similarity rating task to investigate the processing of anterior and posterior coronal sibilants [s] and [ʃ] in three languages in which the two sounds participate in different types of relationships: (i) English, in which [s] and [ʃ] occur in the same environment (e.g. *see* versus *she*); (ii) Korean, in which [s] and [ʃ] are in complementary distribution and participate in regular and productive morphological alternations; and (iii) Mandarin, in which [s] and [ʃ] are in complementary distribution but do not participate in morphological alternations due to its phonotactic restrictions. The results showed that both English and Mandarin speakers rated the anterior versus posterior coronal sibilants as more different than Korean speakers did, suggesting that the Mandarin speakers, who have access only to distributional evidence, are less likely to treat the two sounds as members of a single category than the Korean speakers, who are exposed to evidence from both distribution and morphological alternation.

Key words: alternation, coronal sibilants, distribution, phonological relationships, similarity judgments

1. Introduction

Most phonologists acknowledge the need to distinguish sound differences (or feature differences) that are contrastive from those that are not (e.g. Avery et al. 2008; Steriade 2007). The traditional definition of contrast has relied heavily on the distribution of two given sounds. If two sounds occur in the same environment such that substituting one for the other can signal a lexical difference, then the sounds are considered to be in contrast, belonging to discrete phoneme categories (e.g. Bloch 1948, 1950; Harris 1951; Moulton 1962; Swadesh 1934; Trubetzkoy 1969). The classic test for contrast is the minimal pair test; a minimal pair (e.g. English [si] 'see' versus [ʃi] 'she') consists of two forms with distinct meanings that are phonologically identical except for a single segment (Trubetzkoy 1969). On the other hand, two phonetically similar sounds are considered to be non-contrastive allophones of the same phoneme if they occur in complementary

* I would like to thank Ellen Broselow, José Elias-Ulloa, Kathleen Curie Hall, Marie Huffman, Arthur Samuel, and two anonymous reviewers for their valuable comments and suggestions. I would also like to thank Jiwon Hwang, Hijo Kang, Yuwen Lai, and Yunju Suh for their help with designing and running the experiment. Any remaining errors are mine. This work was supported in part by the Chiang Ching-Kuo Foundation Dissertation Fellowship to the author and by NSF grant BCS-07460227 to Ellen Broselow, Marie Huffman and Nancy Squires.

distribution, where the choice of one sound versus the other is predictable from the environment (Trubetzkoy 1969:46).¹ However, as Hall (2009) has demonstrated, predictability from context is not necessarily a binary proposition. For example, positional neutralizations (e.g. German final devoicing; Charles-Luce 1985) and stratum-specific contrasts (e.g. Japanese native versus foreign strata; Itô & Mester 2001), in which the distributions of two sounds are predictable in some environments but not in others, are often found in human language.

Distributional restrictions may serve as static constraints on lexical items (e.g. the distribution of Korean [s] and [ç] in [se] ‘bird’ and [çi] ‘poem’, where [ç] is only found before the high-front vowel /i/); however, when the choice of variant is dependent on context, we frequently also find morphological alternations: morpho-syntactic processes (e.g. affixation, compounding) that change the phonological context in which a sound appears and may therefore cause a single morpheme to exhibit different variants in different contexts (e.g. the same morpheme /nas-/ realized as [nas-] in [nas-e] but as [naç-] in [naç-i]). Therefore, alternation is sometimes treated as a criterion for determining the relationships among sounds (e.g. Anderson 1985; Baudouin de Courtenay 1972; Silverman 2006). For example, the last sound in the American English morpheme *ride* [ɹaɪd] alternates with [r] when the presence of the suffix *-er* produces a tapping environment (i.e. intervocalic position preceding an unstressed syllable): *rider* [ɹaɪrə].

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).² Morphological alternation has sometimes even been argued to be irrelevant to phonological analyses (e.g. Hockett 1942; Trager 1934). For example, Trager argued that alternation ‘does not properly concern us in a purely phonemic study’ (1934:340). Silverman (2006), on the other hand, argues that ‘learning allophonic relations is dependent upon learning allomorphic relations’ (2006:26) and that ‘the only way sounds can be allophonically related is if they alternate with each other’ (2006:88).

Different phonological models offer different understandings of how these two criteria, distributional predictability and alternation, help determine sound relationships. Phonological approaches that assume economy in phoneme analysis, such as the traditional structuralist approach (e.g. Hockett 1942) and the SPE-type generative approach (e.g. Chomsky & Halle 1968; Clements 2003), hold that as much predictable information as possible should be removed from underlying representations in order to minimize the number of phoneme contrasts/contrastive features in a language. This assumption leads to the inevitable outcome that surface sounds in complementary distribution are derived from the same underlying sound. In output-based OPTIMALITY THEORY (OT; Prince & Smolensky 2004), on the other hand, economy is generally assumed to play a much more limited role. Most researchers in OT assume RICHNESS OF THE BASE:

¹ As one of the anonymous reviewers pointed out, not all allophony is between phonetically similar sounds. For example, allophonic relations between a particular sound and [ø] are a challenge to the phonetic similarity condition.

² Other criteria commonly used to determine phonological contrastiveness 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.

(1) 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, distributional restrictions are regulated by markedness constraints and do not pose a problem at the level of underlying forms. Most researchers in this camp also assume LEXICON OPTIMIZATION, which permits (and requires) two sounds to be derived from a single sound only when they are found in alternating forms of the same morpheme (Inkelas 1995; Yip 1996). The role of morphological alternation is thus made explicit in the OT framework, while the role of distributional predictability in assigning sounds to underlying representations is diminished.

The above discussion makes it clear that predictable distribution, the traditional factor employed in determining contrast, is not without its challenges, particularly since distribution is not a simple binary notion (Hall 2009). Although in some approaches morphological alternation is considered to be merely a reinforcement of distributional predictability (e.g. Hockett 1942; Trager 1934), in other approaches it is seen as an explicit requirement for grouping surface sounds into the same phoneme category (e.g. Inkelas 1995; Silverman 2006).

To determine the contributions of these two factors, distributional predictability and morphological alternation, in leading speakers to group sounds as members of the same phoneme category, this study investigates the processing of anterior and posterior coronal sibilants [s] and [ʃ] in three languages (Mandarin, English, and Korean) in which the two sounds participate in different types of relationships. In Mandarin, the choice of [s] or [ʃ] is largely predictable from the environment: posterior [ʃ] occurs before high-front vowels [i, y] and glides [j, ɥ], and [s] occurs elsewhere (e.g. Duanmu 2007; Lin 2007). However, despite this predictable distribution, the two sounds do not participate in morphological alternations, due to Mandarin's lack of affixation and its stringent restrictions on possible syllable structures. In other words, there is no morphological evidence (e.g. an alternation of [s] and [ʃ] within variants of the same morpheme) to support the relatedness of the two sounds in Mandarin. Due to the conflicting evidence, the relationship between [s] and [ʃ] in Mandarin has been a matter of long-standing controversy (Chao 1931; Cheng 1968, 1973; Chiang 1992; Duanmu 2007; Hartman 1944; Lin 1982; Wan 2010; Wu 1994; Yip 1996; and see a more recent review in Lu 2014). The distributional predictability of these sounds has led some researchers to argue that [s] and [ʃ] should be considered variants of the same category, while the lack of alternations has led others to argue that the two sounds should be treated as members of separate categories. Mandarin [s] and [ʃ] thus provide a good test case to tease apart the contributions of distribution and alternation.³ In this study, I compare Mandarin speakers' treatment of these sounds

³ Though recent studies have commonly taken the view that the two sibilants under consideration, [s] and [ʃ], belong to the same phoneme underlyingly (Duanmu 2007; Lin 2007; Wan 2010), note that the posterior [ʃ] is in complementary distribution not only with the anterior sibilant [s], but also with velar [x]. As an anonymous reviewer has pointed out, no consensus on the allophonic/phonemic overlap among these three sounds has yet been reached. Since the present research considers whether distribution forces speakers to group two sounds into a single phoneme category, either comparison, that is [s] ~ [ʃ] or [ʃ] ~ [x], should be a valid test case. This is similar to the English [d/r] case presented in Boomershine et al. (2008), in which [r] is in complementary distribution with multiple sounds—[d] (*ride* [d] versus *rider* [r]) and [t] (*write* [t] versus *writer* [r]); still more similar ratings were observed for [d/r].

with the treatment of parallel sounds by speakers of two other languages. English differs from Mandarin in that the distribution of anterior and posterior coronal sibilants is not predictable from the environment (e.g. *so* versus *show*; *save* versus *shave*; *sea* versus *she*; *sour* versus *shower*; *lass* versus *lash*), and the difference between the sounds may be used to signal lexical differences. Korean, on the other hand, shares with Mandarin the predictable distribution of [s] and [ʃ], but differs from Mandarin in that these sounds also participate in regular and productive morphological alternations.

The present study investigates Mandarin, Korean, and English speakers' perceptions of anterior and posterior coronal sibilants. The method adopted here has previously been used to demonstrate that variants of the same phoneme are perceived as more similar than sounds belonging to separate phonemes. If distribution alone is sufficient to allow learners to assign two sounds to a single category, we should find that Mandarin and Korean speakers rate the anterior and posterior sibilants as more similar than English speakers, for whom the two sounds belong to different categories based on this criterion. On the other hand, if it is necessary for two sounds to alternate in order for speakers to analyze them as members of a single phoneme category, we would expect Mandarin speakers and English speakers to pattern like one another and unlike Korean speakers, for whom the two sounds regularly alternate. The results of the experiment showed that Mandarin speakers' ratings of the anterior and posterior sibilants differ from those of Korean speakers, and pattern overall with those of English speakers. These results suggest that Mandarin speakers, who have access only to distributional evidence, are less likely to treat the anterior and posterior coronal sibilants as members of a single category than Korean speakers, who are exposed to evidence from both distribution and morphological alternation.

The structure of the paper is as follows: §2 presents background on the languages investigated and the experimental probe used in this study; §3 describes the experimental methodology and reports the results; §4 provides a discussion of the results; and §5 summarizes the results and concludes.

2. Background

2.1 Languages

To tease apart the relative contributions of distribution and alternation in motivating speakers to assign sounds to phonological categories, this study compares the behavior of three groups: Mandarin speakers, for whom the anterior and posterior coronal sibilants are in complementary distribution but do not participate in alternations; English speakers, for whom these sounds are not predictable from the phonological environment; and Korean speakers, for whom these sounds are in complementary distribution and alternate regularly.

2.1.1 English

The English anterior sibilant [s] and posterior sibilant [ʃ] may occur in identical contexts, giving rise to minimal pairs such as *sea* [si] versus *she* [ʃi], *sock* [sək] versus *shock* [ʃək], *sue* [su]

versus *shoe* [ʃu], and *lease* [lis] versus *leash* [liʃ]. These sounds 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* [ɒpriɛs] ~ *oppression* [ɒpriɛʃən]; *press* [priɛs] ~ *pressure* [priɛʃə]).⁴ 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.

As for morphological alternation of the 167 vowel-initial derivational suffixes listed on Wiktionary,⁵ only four suffixes trigger palatalization (i.e. *-ial*, *-ion*, *-ious*, *-ure*), and only two of those provide the pre-palatal context (*-ial* and *-ious*). Thus, distributional evidence in English supports the view that [s] and [ʃ] constitute separate categories, while evidence for grouping them together is weak.

2.1.2 Korean

In Korean, both distribution and alternation point to the analysis of the anterior sibilant [s] and the posterior sibilant [ɕ] as members of a single category. The two sounds occur in distinct environments, as illustrated in (2): [ɕ] occurs before the high-front vowel [i] and glide [j] (mainly in loanwords (a–f)), and [s] occurs elsewhere (g–i) (e.g. Iverson & Lee 2006; Kim 2009; Sohn 1999).

(2) Complementary distribution of Korean [s] and [ɕ]

- | | | |
|----|------------------------|-----------|
| a. | [ɕi] | ‘poem’ |
| b. | [ɕikan] | ‘time’ |
| c. | [ɕjamp ^h u] | ‘shampoo’ |
| d. | [ɕjap] | ‘shop’ |
| e. | [ɕjup ^h ʌ] | ‘super’ |
| f. | [ɕjo] | ‘show’ |
| g. | [sal] | ‘flesh’ |
| h. | [sul] | ‘alcohol’ |
| i. | [se] | ‘bird’ |

Korean [s] and [ɕ] also alternate before different vowel suffixes, as shown in (3). Before the locative suffix *-e*, [s] occurs; before the nominative suffix *-i*, the same morpheme is realized with [ɕ].

⁴ As a reviewer pointed out, the alternation between [s] and [ʃ] in some dialects is largely regular in casual speech, and the optionality arises primarily across different registers. Based on the optionality across dialects and registers, the English case is treated here as a case in which alternation plays a limited role.

⁵ See http://en.wiktionary.org/wiki/Appendix:English_suffixes.

(3) Morphological alternation of [s] and [ɕ] in Korean

a. /nas/	[nas-e]	‘sickle-locative’
	[naɕ-i]	‘sickle-nominative’
b. /kos/	[kos-e]	‘place-locative’
	[koɕ-i]	‘place-nominative’
c. /pus/	[pus-e]	‘writing brush-locative’
	[puɕ-i]	‘writing brush-nominative’

Unlike the anterior and posterior coronal sibilants in English, the two sounds in Korean occur in predictable environments, and participate in productive morphological alternations.

2.1.3 Mandarin

The distribution of the anterior sibilant [s] and posterior sibilant [ɕ] in Mandarin is similar to the distribution of these sounds in Korean: [ɕ] occurs before high-front vowels [i]/[y] and glides [j]/[ɥ], and [s] occurs elsewhere (e.g. Duanmu 2007; Lin 1989, 2007), as shown in (4).

(4) Complementary distribution of Mandarin [s] and [ɕ]

a.	[ɕi3]	‘wash’
b.	[ɕja1]	‘blind’
c.	[ɕjo1]	‘rest’
d.	[ɕje4]	‘crab’
e.	[sa3]	‘spread’
f.	[su1]	‘crispy’
g.	[so1]	‘gather’

However, due to morphological and phonotactic restrictions, these sounds never display alternations. Only dental nasals [n] and velar nasals [ŋ] can occur in coda position, and codas do not re-syllabify to a following onsetless syllable. The sibilants in question do not occur in coda position. Mandarin anterior and posterior coronal sibilants thus provide a good comparison with the equivalent sounds in the other two languages.

Note that the anterior and posterior coronal sibilants in these three languages have different acoustic details (e.g. Beckman & Pierrehumbert 2000; Ladefoged & Maddieson 1996; Li et al. 2007; Suh 2009). For example, the English [s] has a higher centroid frequency than the [s] found in Mandarin and Korean (Beckman & Pierrehumbert 2000; Li et al. 2007), and the English [ʃ] is usually produced with lip protrusion while the Mandarin and Korean [ɕ] is not; furthermore, [ʃ] and [ɕ] differ in the degree of raising of the front of the tongue (Ladefoged & Maddieson 1996:153). In a cross-linguistic comparison like the one undertaken in the current study, such individual language differences are inevitable. The current study employed a perceptual experiment designed to maximally bypass the unavoidable confound caused by these acoustic differences (see §2.2 and §3.1.2).

The different types of relationships between the anterior and posterior coronal sibilants in these three languages are summarized in (5). The parentheses around the English checkmark indicate that the morphological alternation of the two sounds is limited in English.

(5) Languages

	Mandarin	Korean	English
Distribution	√	√	
Alternation		√	(√)

The following section provides a review of the previous literature on the experimental probe used in this study.

2.2 Similarity ratings

Previous studies have shown that phonological relationships have an impact on how individual language users perceive sounds (e.g. Beckman & Pierrehumbert 2000; Best 1991; Boomershine et al. 2008; Dupoux et al. 1999; Johnson & Babel 2010; Werker & Lalonde 1988). The present study employed a previously established experimental method—similarity rating—to investigate the perception of the two coronal sibilants by Mandarin, Korean, and English speakers, in whose languages the two sounds participate in different phonological relationships defined by the two criteria, distributional predictability and morphological alternation.

In similarity rating tasks, speakers tend to rate sounds that exist in an allophonic relationship as more similar than sounds that belong to separate phonemes (Babel & Johnson 2010; Boomershine et al. 2008; Harnsberger 2001). For example, Boomershine et al. (2008) tested native English and Spanish speakers' similarity judgments of [ð], [d], and [r] in different vocalic contexts using an AX paradigm (e.g. [ada]–[ara], [idi]–[iði]). [ð] and [d] are contrastive phonemes in English (e.g. *they* [ðeɪ] versus *day* [deɪ]) but are allophonic variants in Spanish, where intervocalic voiced stops are spirantized following a continuant (e.g. [d] *onde* 'where' but *de* [ð] *onde* 'from where'). Conversely, [d] and [r] are contrastive in Spanish (e.g. [kaða] 'each', [kara] 'face') but are allophonic variants in American English, due to a process whereby [d] (and [t]) becomes a tap intervocalically preceding an unstressed vowel (e.g. *ride* [raɪd] but *rider* [raɪrə]). The phonological relationships among the three sounds are shown in (6).

(6) Phonological grouping of [ð], [d], and [r] in English and Spanish

English	[ð]	[d], [r]
Spanish	[ð], [d]	[r]

Boomershine et al. (2008) asked participants to rate the similarity of a pair of sounds taken from the VCV sequences [ada], [ara], [aða], [idi], [iri], [iði], [udu], [uru], and [uðu]. The vocalic 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 from 1 to 5, where 1 indicated 'very similar' and 5 indicated 'very different'. The results showed a clear native language effect. On the rating task for [r/ð], two sounds that are contrastive in both English and Spanish, the scores for the two language groups converged. However, for the other two pairs, the English speakers rated [d] and [r] as most similar, while the Spanish speakers rated [ð] and [d] as most similar. In each case, the rating pattern

corresponded to the phonological relationship between each pair of sounds in the language under investigation.

Johnson & Babel (2010) tested native English and Dutch speakers' similarity judgments for the fricatives [f, θ, s, ʃ, x, h] using the same methodology as Boomershteyn et al. (2008). The fricative phoneme inventories for Dutch and English are listed in (7).

(7) Voiceless fricative phonemic inventories of Dutch and English (Johnson & Babel 2010)

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

Two factors were examined in this study. First, the effect of allophony was tested. 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 that it derives from an allophonic rule that palatalizes /s/ before /j/; thus, [ʃ] is analyzed as a variant of /s/ before the high-front glide [j] in Dutch (Booij 1999).⁶ Based on these facts, Johnson & Babel (2010) predicted that English listeners should demonstrate a clearer distinction between [s] and [ʃ] than Dutch speakers. Second, the effect of phoneme inventory was tested. Dutch lacks the fricative [θ]. In perceiving this non-native sound, it was predicted that Dutch speakers should have less sensitivity than English speakers towards the distinction between [θ] and the other fricatives, manifested as more similar ratings. Both effects were confirmed by the results. The perceived differences between [s] and [ʃ] and between [θ] and [s]/[ʃ] reported by the Dutch speakers were significantly smaller than those reported by the English listeners. The results suggest: (i) that allophony has an effect on how native speakers perceive sounds, and (ii) that speakers are more likely to assimilate non-native sounds to perceptually similar native sound categories.

In a follow-up experiment, Johnson & Babel (2010) tested the lower-level auditory processing of these same fricatives by English and Dutch listeners. This experiment modified the methodology of the original experiment to include a much shorter inter-stimulus interval (ISI) of 100 ms between A and X, and a 500 ms reaction time target. In this speeded-up task, they found no consistent effect of language of the listener, suggesting that the native phonological knowledge was bypassed. Instead, the speeded-up task tapped into a lower level of processing. The findings of this follow-up experiment indicate that the differences found in Johnson and Babel's original experiment were not due to raw acoustic similarities, and further strengthen the view that native phonological knowledge does influence speakers' perception when high-level processing is encouraged.

⁶ The allophonic rule of palatalization is productive and also involves coronal obstruents and nasals /s, z, t, n/ before /j/ in the native grammar (Booij 1999). McCarthy (2005) proposes that language learners may analyze the non-alternating sounds—non-alternating [ʃ]s in the current case—as /s/ by taking a free ride on the alternating /s/ → [ʃ] mapping.

Note that the ‘allophonic’ cases presented here (i.e. English [d/r], Spanish [d/ð], Dutch [s/f]) are both distributionally predictable and morphologically alternating. Thus, none of the studies discussed so far attempted to tease apart the specific roles of these two factors in phoneme categorization. The present study not only tested further cases where distribution and alternation exist side by side ([s]/[ç] in Korean), but also enriched the findings of previous studies by investigating circumstances in which each factor occurs independently (predictable distribution of [s]/[ç] in Mandarin and alternation of [s]/[f] in English).⁷

The next section describes the methodology of the experiment and reports the results.

3. Experiment

The goal of this study was to compare how English, Korean, and Mandarin listeners rate the perceptual difference between anterior and posterior coronal sibilants in order to tease apart the roles of distribution and alternation in processing sounds as variants of the same phoneme. The results were collected using a similarity rating task. If distribution alone defines the phonological relationship between two sounds, we should expect the Mandarin listeners’ ratings to be similar to those of Korean listeners and different from those of English listeners. If distribution is not a sufficient criterion for phonemic discrimination and the presence of morphological alternation is also necessary in order for listeners to group sounds as variants of the same category, we should expect the ratings of Mandarin listeners and English listeners to pattern with each other and against the ratings of Korean listeners.

3.1 Methodology

3.1.1 Participants

Twenty participants from each of the three language groups were recruited. Participants in the English group (11 male, 9 female, aged 18–22, monolingual native English speakers) and in the Korean group (6 male, 14 female, aged 18–38, native Korean speakers from South Korea) were recruited at Stony Brook University, New York, and received course credit or payment for their participation.⁸ To estimate the possible influence of English, the Korean participants, all of whom had received high-school education in South Korea before coming to Stony Brook, were asked to rate their English ability (Boomershine et al. 2008); the average rating was 4.65 on a seven-point scale. Participants in the Mandarin group (4 male, 16 female, aged 20–22, native speakers of Taiwanese Mandarin) were recruited at National Chiao Tung University in Taiwan for course credit or payment. The average self-reported English proficiency of the Taiwanese participants was 4.6 on a seven-point scale. None of the participants reported any hearing deficiencies.

⁷ Note that the morphological alternation in English, unlike the one in Korean, is restricted to a limited set of suffixes.

⁸ Of the 20 English participants, nine reported that they had taken Spanish, two had taken French, and two had taken Italian in high school, but none had reached a proficient level. The dialectal backgrounds of the English and Korean participants were not controlled.

3.1.2 Design and materials

The experiment employed an AX similarity rating task similar to the ones used in Boomershine et al. (2008) and Johnson & Babel (2010). The materials consisted of the target sibilants [s, ʃ, f], along with two other fricatives [f, h] as controls, embedded in three vocalic contexts [a_a], [i_i], and [u_u]. Materials consisted of two tokens of each of the following VCV sequences: [asa][aça] [aʃa][afa][aha], [isi][içi][ifi][ihi], and [usu][uçu][ufu][uhu]. 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 posterior sibilant [ç] natively 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, using a Marantz digital recorder PMD660 (Marantz, Kanagawa, Japan) and Shure SM 48 microphone (Shure Incorporated, Niles, IL, USA), at a sampling rate of 44,100 Hz. One instance of each VCV was selected as a test item so that the tokens were approximately matched on pitch and duration. Example (8) shows the average pitch of the first and second vowels of the selected stimuli, and (9) shows the vowel and fricative durations of the selected stimuli. 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).

(8) Pitch in Hz of the first and second vowels

	V1	V2
aça	118	117
afa	120	118
aha	119	117
aʃa	114	116
asa	114	116
içi	113	115
ifi	114	114
ihi	113	114
ifı	114	114
isi	114	112
uçu	118	118
ufu	118	119
uhu	116	118
ufu	116	118
usu	117	117
	V1 averaged pitch: 115.87 Hz	V2 averaged pitch: 116.2 Hz
	Standard deviation: 2.33 Hz	Standard deviation: 2.01 Hz

- (9) Durations in ms of the first vowel, the fricative, the second vowel, and the total duration of the stimulus⁹

	V1	Fric	V2	Total
aca	225	198	330	753
afa	301	142	333	775
aha	277	128	313	717
asa	266	162	337	765
aʃa	232	168	294	694
ici	255	201	320	776
ifi	262	152	329	743
ihi	278	137	305	720
ifi	213	216	305	734
isi	243	182	309	734
ucu	194	213	315	722
ufu	226	169	291	685
uhu	231	152	288	671
ufu	222	227	263	712
usu	203	196	297	695
				Averaged duration: 726.4 ms
				Standard deviation: 32.28 ms

Note that the posterior sibilant [ɕ] does not exist in the English consonant inventory, [f] does not exist in Korean, and [ʃ] does not exist in either Korean or Mandarin. Also note that some of the stimuli contained sequences that violated the phonotactics of one or more of the languages: *[si] and *[ɕu] in Korean/Mandarin; *[fi] and *[hi] in Mandarin. Since the task involved the perception of non-native sounds, this experiment followed Johnson & Babel (2010) in employing an ISI of 1000 ms between A and X in order to ensure a high level of processing and assimilation of these non-native sounds to native categories. With this longer ISI, a lower (acoustic) level of processing could be maximally bypassed (Pisoni 1973; Werker & Logan 1985) to reach a phonological level of processing.

The experiment employed a three-factorial design with one between-subject factor (Language) and two within-subject factors (Pair, Vowel Context), as shown in (10).

⁹ The frication durations of English [s] and [ʃ] are both reported to be around 178 ms (Jongman et al. 2000) and those of Korean tense [s'] and lax [s] are reported to be around 160 ms and 100 ms, respectively (Cheon & Anderson 2008). The frication durations of the stimuli fall within the range of those in English and Korean.

(10) Similarity rating design

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

The pairings of the five fricatives yielded $5^2 = 25$ possible pairs, including five pairs in which both members were the same ($5^2 - 5$ identical pairs = 20 different pairs). Each different pair was presented once, while the identical pairs were presented twice in order to balance the number of same and different pairs, yielding 30 trials (20 different pairs + 5×2 same pairs = 30) per vowel context (30×3 vowel contexts = 90). Listeners heard each of the AX trials (90 trials) three times in three blocks (90×3 blocks = 270). Participants had a maximum of 5000 ms to respond to a given trial before the next trial started.

3.1.3 Procedure

Participants were presented with written instructions on a computer screen in their native language. The instructions stated that the participant 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, using a computer that was connected to a keyboard with five keys labeled from 1 to 5. E-Prime software (v2.0; Psychological Software Tools, Pittsburgh, PA) was used to ensure that the pairs were presented in a different random order for each participant. All stimuli were presented binaurally over headphones at a comfortable listening level.¹⁰ The participants completed a nine-trial practice session compiled randomly from the test trials, and had the opportunity to ask questions before proceeding to the experiment. The experiment lasted approximately 20 minutes.

3.2 Results

We expect to find the greatest difference in rating between the anterior sibilant [s] and the posterior [ɕ/ʃ] among the English listeners, for whom the two sounds are not predictable from the environment and participate only in limited alternation; we expect to find more similar ratings between the two sounds for the Korean listeners, for whom the two sounds are predictable and alternate regularly. Of particular interest here are the Mandarin listeners’ ratings. If distribution alone

¹⁰ The headphones used for the Mandarin participants (AKG K240MKII) were different from the ones used for the Korean and English participants (Philips HSP805) due to the different locations in which the experiments were conducted. Given this fact, there is a possibility that language-specific effects could be found.

defines the phonological relationship between the two sounds, we should expect the Mandarin listeners' ratings to be similar to those of the Korean listeners. If alternation contributes to the determination of phonological relationships, we should expect the ratings of the Mandarin listeners to be more different from those of the Korean listeners.

The rating scores for each participant were normalized into *z*-scores (by taking the difference between the individual score and the mean divided by standard deviation) to compensate for differences among participants in using the five-point scale (Boomershine et al. 2008). The standardized scores were centered around 0, with scores above 0 indicating 'more different' and scores below 0 indicating 'more similar'. The normalized results are illustrated in (11).¹¹ The *x*-axis represents the different sound pairs and the *y*-axis represents the normalized *z*-scores.

From (11), we can see that, with the exception of the target pairs indicated by the arrows, ratings from the three language groups were comparable.¹² The differences lay in the target pairs

¹¹ The means of the raw scores and *z*-scores are reported here with standard deviations in parentheses. Note that the raw scores for the English group are lower overall than those for the Korean group, and those for the Korean group are in turn lower overall than those for the Mandarin group. As one of the reviewers pointed out, this could be due to the familiarity of the stimuli since the stimulus recording was done by a Mandarin native speaker.

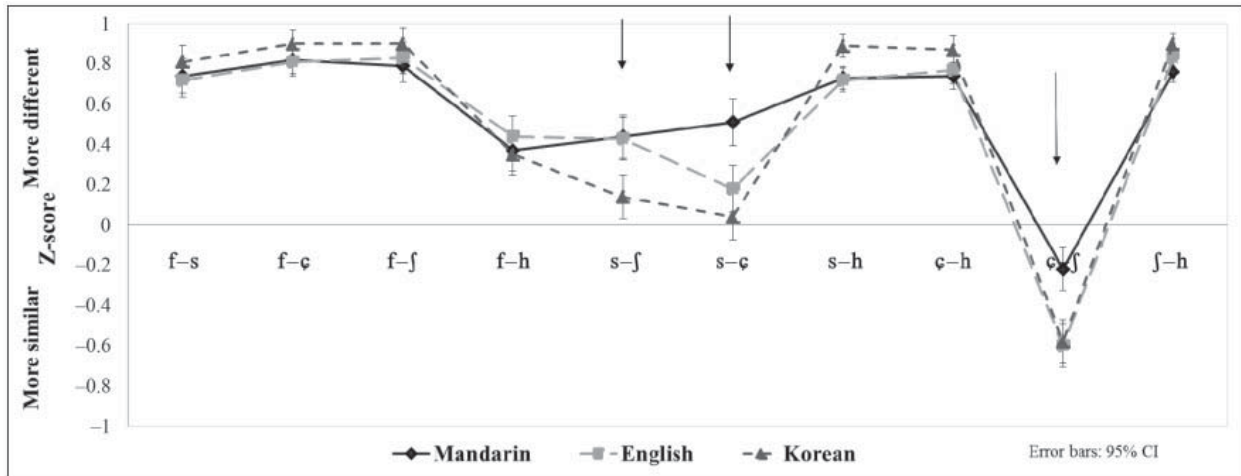
Pair		[f-s]	[f-ɕ]	[f-f]	[f-h]	[s-f]	[s-ɕ]	[s-h]	[ɕ-h]	[ɕ-f]	[f-h]
English	Raw	3.35	3.46	3.51	3.01	2.96	2.62	3.36	3.40	1.65	3.49
	scores	(1.09)	(1.07)	(1.11)	(1.19)	(1.04)	(1.08)	(1.09)	(1.09)	(.86)	(1.05)
	Z-scores	.72 (.66)	.81 (.68)	.83 (.68)	.44 (.79)	.43 (.68)	.18 (.77)	.72 (.66)	.77 (.72)	-.60 (.64)	.83 (.65)
Korean	Raw	4.16	4.30	4.31	3.42	3.09	2.94	4.28	4.27	4.27	4.28
	scores	(1.01)	(.95)	(.98)	(1.42)	(1.24)	(1.30)	(1.03)	(1.03)	(1.03)	(1.00)
	Z-scores	.81 (.53)	.90 (.49)	.90 (.51)	.35 (.80)	.14 (.65)	.04 (.69)	.89 (.55)	.87 (.54)	-.58 (.56)	.90 (.58)
Mandarin	Raw	4.30	4.39	4.33	3.69	3.80	3.91	4.24	4.25	2.73	4.28
	scores	(.98)	(.91)	(.95)	(1.26)	(1.24)	(1.21)	(.98)	(1.06)	(1.43)	(.99)
	Z-scores	.74 (.48)	.82 (.48)	.79 (.52)	.37 (.67)	.44 (.64)	.51 (.64)	.73 (.48)	.74 (.60)	-.22 (.81)	.76 (.53)

¹² The results for the identical pairs have been removed from the main text for the sake of space. The numbers are reported here with standard deviations in parentheses. As the numbers show, these pairs were rated as very similar for all three groups of participants.

Pair		[f-f]	[s-s]	[ɕ-ɕ]	[f-f]	[h-h]
English	Raw scores	1.04 (.30)	1.09 (.43)	1.14 (.49)	1.10 (.44)	1.08 (.43)
	Z-scores	-1.05 (.23)	-1.02 (.34)	-.98 (.37)	-1.01 (.31)	-1.04 (.29)
Korean	Raw scores	1.11 (.46)	1.31 (.71)	1.26 (.68)	1.22 (.60)	1.13 (.56)
	Z-scores	-1.11 (.31)	-.98 (.41)	-1.01 (.44)	-1.03 (.39)	-1.09 (.44)
Mandarin	Raw scores	1.15 (.51)	1.23 (.63)	1.27 (.65)	1.31 (.70)	1.31 (.82)
	Z-scores	-1.21 (.33)	-1.16 (.36)	-1.14 (.35)	-1.10 (.46)	-1.11 (.45)

[s-ʃ], [s-ç], and [ç-ʃ]. An omnibus ANOVA (between-subject variable [Language] \times within-subject variable [Vowel, Pair]) showed a main effect of Vowel ($F(2, 114) = 89.48, p < .001, \eta_p^2 = .61$) and of Pair ($F(9, 117) = 272.79, p < .001, \eta_p^2 = .83$), and an interaction between Language \times Pair ($F(18, 117) = 8.63, p < .001, \eta_p^2 = .23$), Language \times Vowel ($F(4, 114) = .77, p < .05, \eta_p^2 = .09$), and Vowel \times Pair ($F(18, 1026) = 43.10, p < .001, \eta_p^2 = .43$). I shall discuss the effect of language in §3.2.1 and the effect of vowel context in §3.2.2.

(11) Similarity rating normalized results



3.2.1 Effect of language

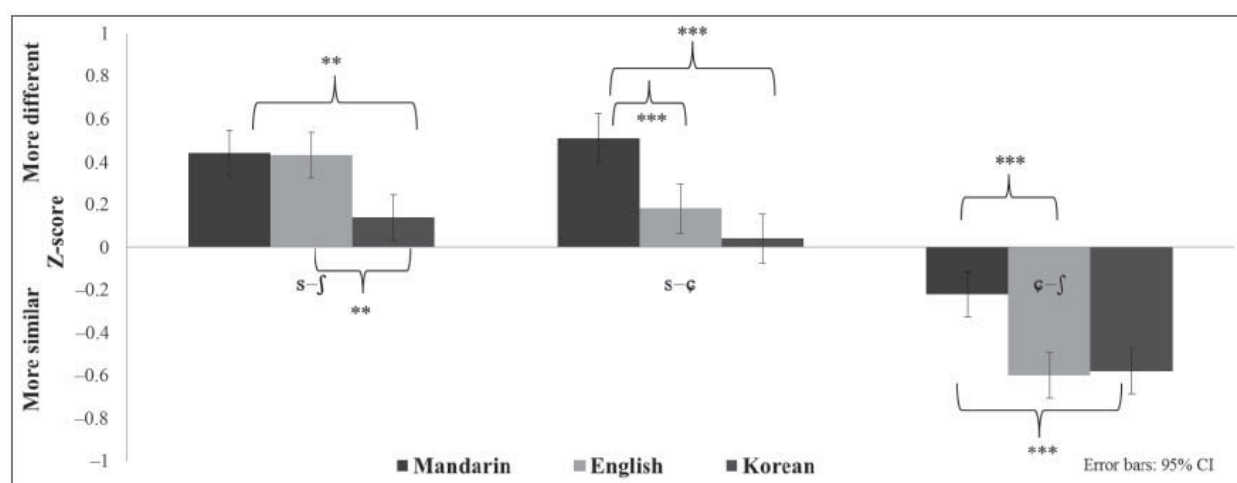
A mixed-factorial analysis of variance (ANOVA) (Language [Mandarin, English, Korean] \times Pair [s, ç, ʃ, f, h]) was performed to interpret the results. The analysis showed a main effect of Language ($F(2, 57) = 7.962, p = .001, \eta_p^2 = .99$) and of Pair ($F(9, 513) = 273.419, p < .001, \eta_p^2 = .83$). Most importantly, there was a significant Pair by Language interaction ($F(18, 513) = 8.647, p < .001, \eta_p^2 = .23$), meaning that the ratings for pairs of sounds were statistically different depending on the native language of the participants. The statistical results are summarized in (12) (*: $p < .05$; **: $p < .01$; ***: $p < .001$).

(12) Summary of similarity rating results

**Language
 ***Pair
 ***Language \times Pair

Of interest here are the ratings of the target pairs [s-ʃ], [s-ç], and [ç-ʃ], as shown in (13). The x-axis represents the three target sound pairs and the y-axis represents the normalized z-scores.

(13) Anterior/posterior coronal sibilants similarity rating results



The [s-f] and [s-c] pairs (the two anterior versus posterior sibilant 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-f] pair ($F(2, 57) = 10.243$, $p < .001$, $\eta_p^2 = .04$). Post-hoc tests showed that the significance came from Mandarin versus Korean, and English versus Korean (both $p < .01$).¹³ The ratings from Mandarin versus English were not significantly different ($p = .991$). That is, the Mandarin and English groups patterned the same for the [s-f] pair, while the Korean group rated these sounds as significantly more similar than the other two groups.

The factor Language yielded a significant effect in the [s-c] pair as well ($F(2, 57) = 17.510$, $p < .001$, $\eta_p^2 = .07$). The significance came from post-hoc tests of Mandarin versus English and Mandarin versus Korean (both $p < .001$). The standardized rating scores produced by the English group were numerically higher than those produced by the Korean group (meaning that [s] and [c] were perceived as more different by the English listeners than by the Korean listeners), although the difference was not significant ($p = .214$). For the [s-c] pair, though the Korean and English groups patterned similarly, this pattern was induced by a certain vowel context (i.e. the [i_i] context; see the discussion in the next section).

As for the ratings for the [c-f] pair (Korean and Mandarin [c]; English [f]), we can see from (13) that listeners from all three languages rated these two posterior sibilants as very similar (all below 0), although the Mandarin listeners' ratings were higher overall (i.e. more different). Subsequent analyses showed that the factor Language was significant in the [c-f] pair ($F(2, 57) = 15.859$, $p < .001$, $\eta_p^2 = .06$). Post-hoc tests showed that the significance came from the Mandarin versus English groups and the Mandarin versus Korean groups (both $p < .001$). There was no significant difference between the English and Korean groups ($p = .967$). A possible explanation for these results will be provided in the next section.

¹³ A Tukey procedure was used for post-hoc tests to control the family-wise error rate over the entire set of pair-wise comparisons.

The statistical results for the target sound pairs are summarized in (14).

(14) Summary of similarity ratings on the anterior/posterior sibilant pairs

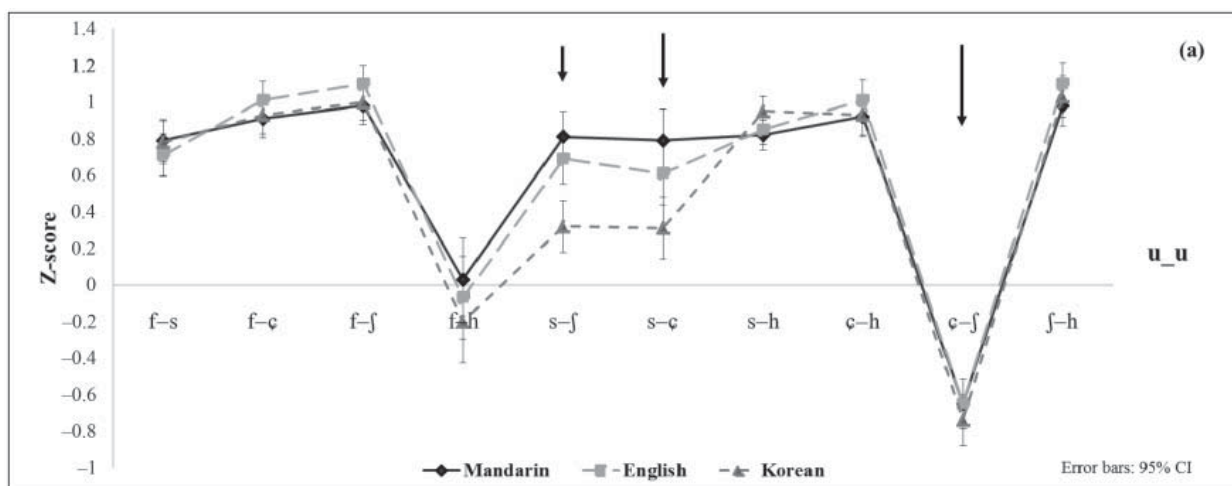
***Simple effect of Language in [s-f]	**Mandarin & Korean **English & Korean Mandarin & English ($p = .991$)
***Simple effect of Language in [s-ɕ]	***Mandarin & English ***Mandarin & Korean English & Korean ($p = .214$)
***Simple effect of Language in [ɕ-f]	***Mandarin & English ***Mandarin & Korean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

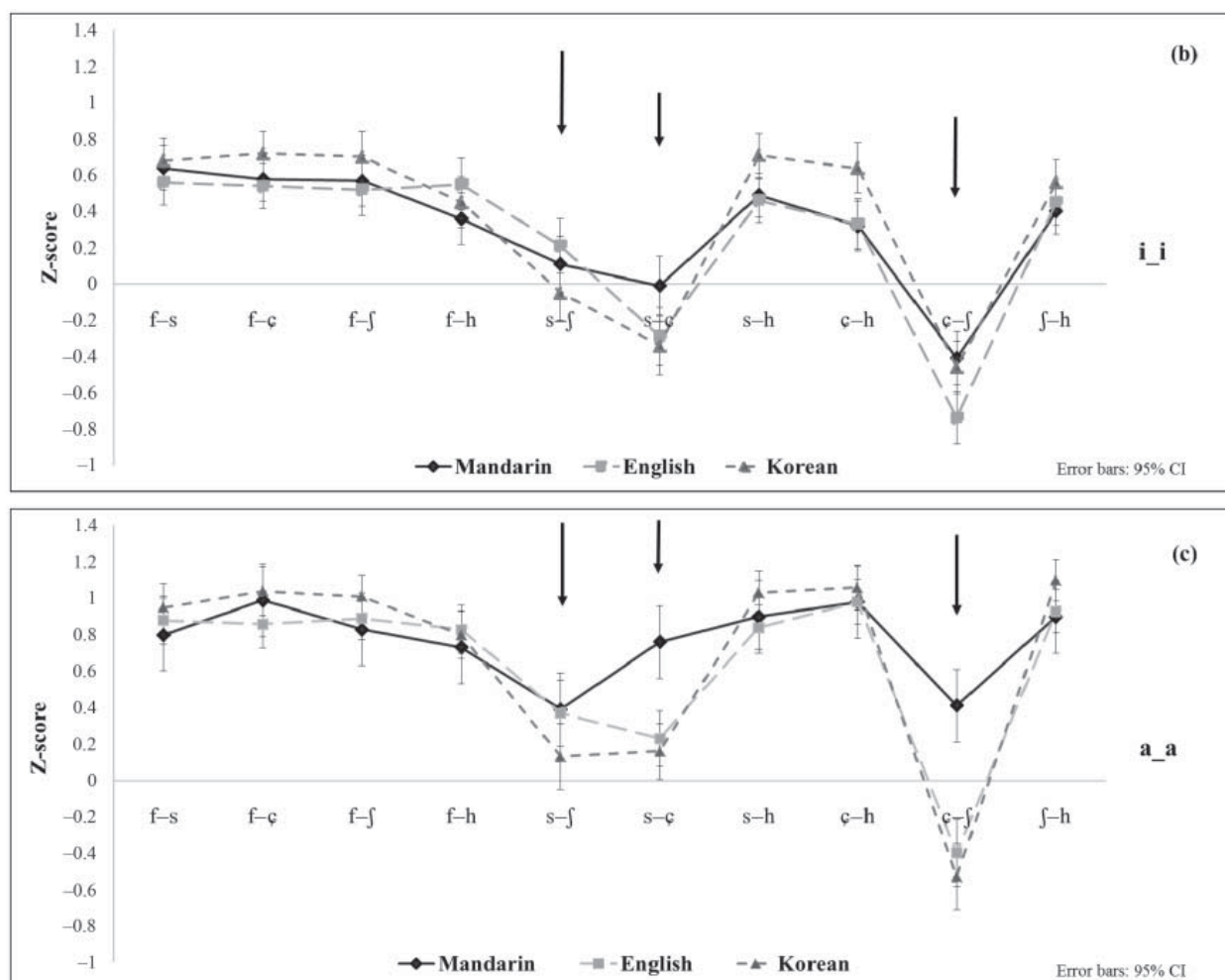
To summarize the overall results: the [s-f] pair was rated as more different by the Mandarin and English groups than by the Korean group; the [s-ɕ] pair was rated as more different by the Mandarin group than by the English and Korean groups; and the [ɕ-f] pair was rated as more different by the Mandarin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

3.2.2 Effect of Vowel Context

The mixed-factorial ANOVA including Vowel Context as a variable (Language [Mandarin, English, Korean] \times Pair [s, ɕ, f, h] \times Vowel Context [a_a], [i_i], [u_u]) yielded some unexpected results, as shown in (15). The x -axis represents the different sound pairs, and the y -axis represents the normalized z -scores. From this figure ((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).

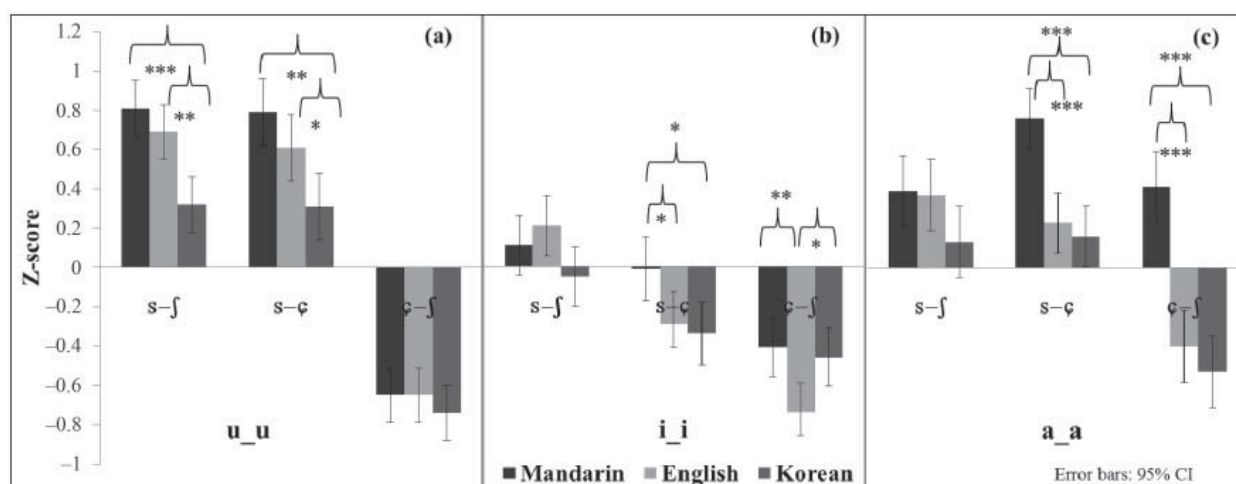
(15) Similarity rating standardized results by Vowel Context





Of interest here are the ratings of the target pairs [s-j], [s-ç], and [ç-j], as shown in (16) ((a): [u_u]; (b): [i_i]; (c): [a_a]).

(16) Anterior/posterior coronal sibilants similarity rating results by Vowel Context



The Korean group showed a general tendency to provide stronger similarity ratings (manifested as lower *z*-scores) for all the target pairs than the other two groups did. The one exception to this tendency occurred in the [c–ʃ] pair. In the [u_u] context (16a), 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 ($F(2, 60) = 13.124, p < .001, \eta_p^2 = .09$), and post-hoc tests showed that the significance came from the difference between the Mandarin and Korean groups ($p < .001$), and between the English and Korean groups ($p < .01$). No significant difference was found between the Mandarin and English groups ($p = .459$). A significant Language effect was found in the [s–c] pair ($F(2, 60) = 17.510, p < .001, \eta_p^2 = .09$). The significance came from the difference between the Mandarin and Korean groups ($p < .01$), and between the English and Korean groups ($p < .05$). No effect was found for the [c–ʃ] pair ($F(2, 60) = .619, p = .542$). The results for the [u_u] context showed that the Mandarin and English listeners treated the anterior sibilant [s] and the posterior sibilants [ʃ/c] as more different from each other than the Korean listeners did.

In the [i_i] context (16b), the Korean listeners again rated the target pairs as more similar than the other two groups did; again, the only exception was the [c–ʃ] pair, which the English listeners rated as most similar. No statistical difference was found among the three groups 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–c] pair ($F(2, 60) = 4.966, p < .05, \eta_p^2 = .05$), with the significance coming from the difference between the Mandarin and English groups, and between the Mandarin and Korean groups (both $p < .05$); the difference in the ratings of the English and Korean groups was not statistically significant ($p = .896$). There was a significant Language effect for the [c–ʃ] pair ($F(2, 60) = 5.810, p < .01, \eta_p^2 = .05$), although the *z*-scores from all three language groups were below 0, indicating that the two sounds were very similar for all the listeners. The significance came from the difference between the Mandarin and English groups ($p < .01$), and between the Korean and English groups ($p < .05$). There was no significant difference between the Mandarin and Korean groups ($p = .898$).

In the [a_a] context (16c), the [s–c] pair was rated as more different by the Mandarin listeners than by the other two groups, and a trend towards greater 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–c] ($F(2, 57) = 18.642, p < .001, \eta_p^2 = .16$) and [c–ʃ] pairs ($F(2, 57) = 5.187, p < .001, \eta_p^2 = .28$). Both of the significant effects came from the difference between the Mandarin and Korean groups, and between the Mandarin and English groups (all $p < .001$). No significant difference was found between the English and Korean groups ([s–c]: $p = .752$; [c–ʃ]: $p = .599$).

Note that in the [a_a] context, the Mandarin speakers rated the [c–ʃ] pair as more different than the other two groups did. This result is surprising, given the fact that, in each language, one of these sounds is absent: [ʃ] does not exist in Mandarin/Korean, and [c] does not exist in English. Why, then, should Mandarin speakers perceive [c–ʃ] as more different than speakers from the other two groups? A possible explanation for this discrepancy is that Mandarin speakers usually perceive and adopt the lip protrusion of the nonexistent posterior sibilant [ʃ] to their native language as a component of the front-rounded vowel [y] or glide [ɥ] (e.g. *Josh* → [tɕ^hjawɕy]; *Michelle* → [miɕɥɛɿ]).¹⁴ Since the design of the current experiment encouraged assimilation of non-native sounds

¹⁴ An anonymous reviewer points out that the loan adaptation of [ʃ] with rounding is Taiwan Mandarin-specific. Mainland Mandarin, for example, does not do this, and thus the rounding might not have to do with the phonology of the language, but with dialectal idiosyncrasies. Since the Mandarin participants were all Taiwan Mandarin speakers, the explanation proposed earlier is still valid.

to native sound categories (see §2.2 and §3.1.2), the more different ratings might be due to the fact that Mandarin participants perceived the rounding on [ʃ] as signaling the front-rounded glide [ɥ]. In other words, the Mandarin listeners might have been comparing the similarity of [aɕa] and [aɕɥa], and as a consequence have perceived the sound pair as more different than the English and Korean listeners did.

The lower similarity ratings given by the Korean and Mandarin groups for the [ɕ-ʃ] pair in the [i_i] context could be explained the same way. Korean speakers, like Mandarin speakers, usually perceive and adopt the rounding of the non-native [ʃ] to their native language as a labial glide [ʷ] (e.g. *she* → [ɕʷi]; *Schick* → [ɕʷikʰi]) (Suh 2009). Thus, the Korean listeners might have been comparing the similarity of [iɕi] and [iɕʷi], and thus rating the sound pair as more different than the English listeners did. Conversely, 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 non-native [ʃ] is perceived as [ɕ] (see (c) and (d) in (2)). Thus, Korean listeners rated the [ɕ-ʃ] pair as very similar. The statistical results are summarized in (17).

(17) 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 ($p = .599$)

4. Discussion

We expected the English listeners to judge the anterior sibilant and posterior sibilant pairs ([s-ç] and [s-f]) as more different than the Korean listeners did, based on the different phonological relationships among the sound pairs in these two languages: in English, although they participate in a limited alternation, [s] and [f] may signal differences in meaning, and the choice of one versus the other is not predictable from the environment; in Korean, on the other hand, the occurrence of the anterior versus posterior sibilants is predictable based on distribution, and the two sounds participate in regular and productive morphological alternations. The overall results suggest that, as has already been reported in many previous studies (e.g. Boomershine et al. 2008; Goto 1971; Johnson & Babel 2010; Werker et al. 1981), the phonological relationship affects the perceived phonetic similarity of two sounds. We found that the Korean listeners rated the anterior and posterior sibilants (both [s-ç] and [s-f]) 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 the same direction for the [a_a] and [i_i] contexts. Furthermore, the results from the English listeners, who rated the [s-f] pair as very different, replicated those of Johnson & Babel (2010).

The similar judgments for the [ç-f] pair among the English group echo the findings reported in Lisker (2001) and McGuire (2007) that English listeners in general categorize the non-native sound [ç] as /f/. Along the same lines, we should expect English listeners' judgments on the [s-ç] and [s-f] pairs to be similar, since [ç] and [f] are perceived as members of the same category (/f/) (Lisker 2001; McGuire 2007). The prediction holds for [a_a] and [u_u]: the ratings of the [s-ç] pair were not significantly different from those of the [s-f] pair in these contexts ([a_a]: $F(1, 19) = 1.514$, $p = .234$; [u_u]: $F(1, 19) = 1.197$, $p = .288$). However, this prediction does not hold for the [i_i] context ($F(1, 19) = 35.374$, $p < .001$, $\eta_p^2 = .15$): English listeners' judgments on the [s-ç] pair in the [i_i] context were significantly different from those on the [s-f] pair. Kathleen Hall (personal communication, 2011) suggests that this pattern, in which English listeners judged [s-ç] as more similar in the [i_i] context than in the other vowel contexts, might be explained as follows: English listeners, encountering the non-native sound [ç] in their native language, might have perceptually assimilated it to the category /f/ in the [u_u] and [a_a] contexts (see Lisker, 2001; McGuire, 2007), but treated [ç] as a positional variant of [s] in pre-palatal position, the [i_i] context. This account could explain why the English and Korean groups appear to pattern together on ratings of the [s-ç] pair, an otherwise surprising result since the anterior and posterior sibilants participate in different phonological groupings in the two languages: the more similar ratings of this pair from the English group were mainly driven by the [i_i] context.

The results from the English and Korean groups suggest that the similarity rating task employed in this study, along with equivalent tasks that have been used in previous studies, does reflect the phonological relationships between sounds in the participants' native languages: English listeners judged the anterior and posterior sibilants as more different than the Korean listeners did. We can now turn to the Mandarin results. If distributional predictability in the absence of morphological alternation is not a sufficiently robust indicator to allow two sounds to be grouped as variants of the same phoneme, then we should expect the Mandarin listeners' ratings of the anterior and posterior sibilants to be more different than those of the Korean speakers. If distributional

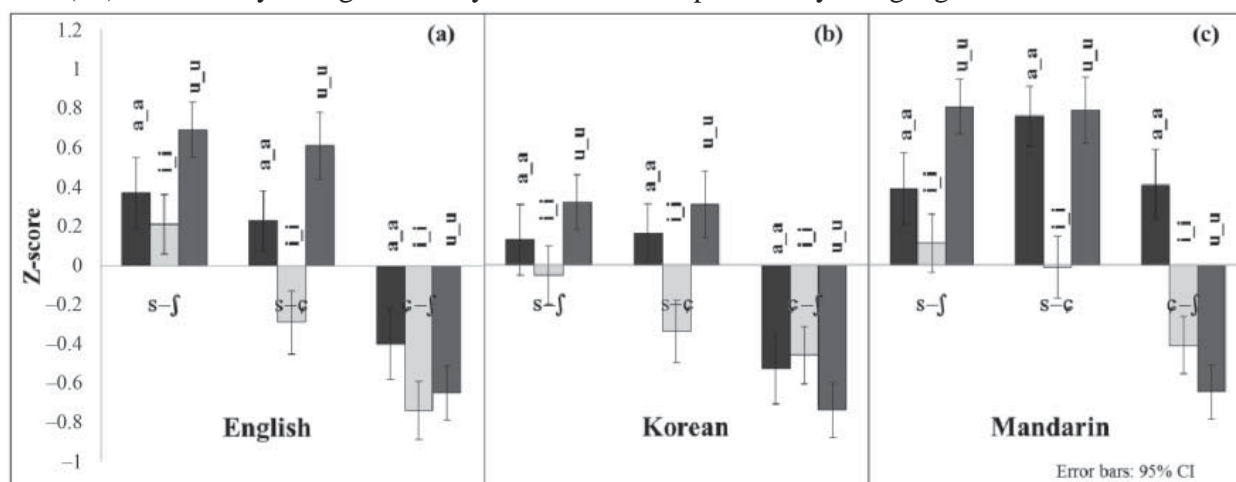
predictability alone is sufficient to group two sounds as variants of the same phoneme, then we should expect the Mandarin speakers' ratings to be comparable to those of the Korean speakers.

We found that the Mandarin listeners rated the anterior and posterior sibilants (both the [s-ç] and [s-ʃ] pairs) 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-ç] pair (whose components exist in complementary distribution, but do not alternate) as significantly more different than did the speakers of Korean, for whom the two sibilants do alternate. This suggests that Mandarin listeners, who are exposed to only distributional evidence, are less likely to group the two sounds as variants of the same category than Korean listeners, who are exposed to both distributional evidence and morphological alternation.

The possibility arises that it is differing levels of English proficiency, rather than a difference in phonological relationships, that causes this Mandarin–Korean discrepancy; perhaps the Mandarin speakers have established a posterior sibilant category as a result of greater exposure to English. However, the Korean group actually rated their English ability slightly higher (4.65/7.00) than the Mandarin group did (4.6/7.00); furthermore, 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 the two sounds in the Korean group as compared to the Mandarin group.

Of interest here are the by-vowel context results, as shown in (18) ((a): English; (b): Korean; (c): Mandarin). The scores above 0 indicate 'more different' and the scores below 0 indicate 'more similar'.

(18) Similarity rating results by Vowel Context paneled by Language



The similarity judgments from the listeners within each language group varied depending on the vowel contexts for different sound pairs. Most importantly, although the Mandarin speakers' similarity judgments patterned overall with those of the English speakers—the English and Mandarin speakers rated the anterior and posterior sibilants as more different than did the Korean speakers—we found that the similarity judgments of the listeners on these sound pairs from all three languages varied according to the vowel contexts.

For all three language groups, the perceived perceptual similarity within the anterior–posterior pairs ([s–ɕ] and [s–ʃ]) increased in the [i_i] context. It is not surprising that this context elicited increased similarity ratings from the Korean (18b) and Mandarin listeners (18c), since in both these languages, only one of the two sounds, [ɕ], can occur before the vowel *i*. The increased similarity judgments from the Korean and Mandarin groups, though significantly more similar from the Korean group than from the Mandarin group, suggest that both *distribution* and *alternation* are relevant. In particular, the fact that these sounds were rated as more similar 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 used in deciding sound membership, we should have seen a reduced phonetic distance between the anterior and posterior sibilants 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 in the [i_i] context was significantly less robust in the Mandarin group than in the Korean group; this finding suggests an effect of alternation. That is, if distribution were the only factor at play in grouping sounds as members of the same category and alternation were irrelevant, we should have seen a similar degree of reduced phonetic distance between the anterior and posterior sibilants in the [i_i] context and in the other vowel contexts in the two language groups.

Interestingly, perceptual similarity also increased for the English listeners in the [i_i] context (18a) (although English speakers still did not rate the sound pair similarity as high as the other two groups did). In other words, vowel context had an effect on similarity rating even in a language in which the two sounds under comparison are contrastive and signal lexical differences in all the given contexts (e.g. *see* versus *she*, *sue* versus *shoe*, *sock* versus *shock*). Contrary to expectations, 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 result is that, although the experiment was conducted in such a way as to maximally bypass the acoustic level of processing, in the pre-palatal context (before the high-front vowel [i]), the place of articulation of the anterior sibilant is produced with some palatalization, and thus the phonetic distance between the anterior and posterior sibilants is reduced. Another explanation for the increased perceptual similarity in English may be that, in connected speech, the anterior sibilant alternates with the posterior sibilant in pre-palatal contexts (e.g. *miss* [mis] ~ *miss you* [mɪʃju]). The perceived phonetic distance between the two sounds in the [i_i] context might be reduced because of this alternation. In other words, the reduced perceptual distance could be due to the effect of phonetic alternation.

One might also wonder whether the increased perceptual similarity reported by the English group in the [i_i] context was due to the morphological alternations discussed in §2.1.1. This possibility is unlikely, however, since the morphological alternations between the two sounds are limited to certain suffixes that do not necessarily provide the pre-palatal context. In other words, these alternations are morphologically conditioned, and do not necessarily depend 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, just as we did in Korean.

An anonymous reviewer expresses concern that the recording of the stimuli by a native Mandarin speaker might be a confound in the relative perceptual prowess of the Mandarin versus English and Korean speakers. In other words, the across-the-board better perception among the target sound pairs in different vowel contexts from the Mandarin participants could be due to the

‘Mandarin-ness’ of the stimuli, instead of the lack of alternation. However, while familiarity of the stimuli might be a factor, the effect should be very marginal since, except for the three target pairs, the rating scores were not significantly different across the three language groups. That is to say, if familiarity were the main cause driving the better perception, we should have seen the same effect for other sound pairs.

The same reviewer suggests another possible interpretation for the better perception from the Mandarin group: due to a phonotactic restriction where [ç] only occurs before [i/j], Mandarin speakers might hear an onglide between [ç] and the following segment in other vocalic contexts, namely [açja] and [uçju]. If that is the case, then the better perception from the Mandarin listeners between [s] and [ç] was not due to the lack of allophonic relation between the two sibilants, but because these speakers were sensitive to the \emptyset versus [j] difference. However, Korean is subject to the same phonotactic restriction. If the better perception from the Mandarin group were due to the \emptyset versus [j] difference, we should have observed the same facilitation in the Korean group.

Another possible confound, pointed out by another reviewer, is the orthographic difference in the representation of [s] and [ç] in Mandarin versus Korean. In Mandarin, both Zhuyin Fuhao and Pinyin spell the two sounds differently, while in Korean, the two sounds are spelled with the same symbol. This is an unavoidable confound unless pre-school-aged children are used as participants to see if the same results are found. This will be left for future study.

To summarize the results, 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 (anterior/posterior sibilants 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 (anterior/posterior sibilants in Korean). We also found that the similarity ratings in the pre-palatal context ([i_i]) were significantly lower for the Korean group than for the Mandarin group, suggesting that distribution reinforced by alternation produces a stronger motivation for the postulation of a unique underlying phoneme than does distribution alone.

5. Conclusion

This paper began by speculating about the sort of evidence that causes native speakers to analyze two sounds as members of a single phoneme category. I investigated how native speakers of English, Korean, and Mandarin rated the similarity of anterior and posterior coronal sibilants, two sounds that participate in different phonological relationships in these three 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 the two sounds are in complementary distribution and participate in productive morphological alternations, rated these sounds as more similar than did the English listeners, in whose language the two sounds are not predictable from the phonological environment. The similarity judgments of the Mandarin group, in whose language the two sounds 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 conclusion that alternation reinforces the mapping of two

sounds to the same category, giving a stronger effect than that of distribution alone. These results also lend support to phonological models in which the principle of economy is not assumed in sound relationships and in which the role of morphological alternation is made explicit (e.g. the assumption of Lexicon Optimization in OT).

Several important questions have emerged from these findings. If both distributional predictability and morphological alternation contribute to the determination of sound membership, this raises the question of why, in the case of Mandarin, the lack of morphological alternation between the two sounds seems to be taking precedence over the predictability of the phonological environments (indicated by the overall more different ratings of the anterior and the posterior sibilants in the Mandarin group versus the Korean group). This is an interesting finding, given that alternation does not seem to account for the English results in the same experiment (indicated by the overall more different ratings in the English group versus the Korean group). The limited morphological alternation discussed above might explain why the unpredictability of the two sounds from the phonological environments seems to be taking precedence in the English case. This speculation leads to a follow-up question: Is there a frequency threshold of a certain factor (e.g. the limited morphological alternation in English versus the productive alternation in Korean) that pushes learners to categorize two sounds one way or another? In other words, how limited is limited and how productive is productive? Directions for future research include the introduction of quantitative measures for each of the criteria used to measure phonological contrastiveness and the weighting of different factors that play a role in determining phonological relationships.

References

- Anderson, Stephen R. 1985. *Phonology in the Twentieth Century: Theories of Rules and Theories of Representations*. Chicago: University of Chicago Press.
- Avery, Peter, B. Elan Dresher & Keren Rice. (eds.) 2008. *Contrast in Phonology: Theory, Perception, Acquisition*. Berlin: Mouton de Gruyter.
- Babel, Molly & Keith Johnson. 2010. Accessing psycho-acoustic perception and language-specific perception with speech sounds. *Laboratory Phonology* 1.1:179–205.
- Baudouin de Courtenay, Jan. 1972. An attempt at a theory of phonetic alternations. *A Baudouin de Courtenay Anthology: The Beginnings of Structural Linguistics*, ed. and translated by Edward Stankiewicz, 144–212. Bloomington: Indiana University Press.
- Beckman, Mary E. & Janet B. Pierrehumbert. 2000. Positions, Probabilities, and Levels of Categorization. Keynote address at the 8th Australian International Conference on Speech Science and Technology (SST-8), December 4–7. Canberra: Australian Speech Science and Technology Association Inc.
- Best, Catherine T. 1991. The emergence of native-language phonological influences in infants: a perceptual assimilation model. *Haskins Laboratories Status Report on Speech Research* SR-107/108:1–30.
- 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. *Glott International* 5.9/10: 341–345.
- Booij, Geert. 1999. *The Phonology of Dutch*. Oxford: Clarendon Press.
- Boomershine, Amanda, Kathleen Currie Hall, Elizabeth Hume & Keith Johnson. 2008. The impact of allophony versus contrast on speech perception. *Contrast in Phonology: Theory, Perception, Acquisition*, ed. by Peter Avery, B. Elan Dresher & Keren Rice, 145–171. Berlin & New York: Mouton de Gruyter.
- 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* 2:320–354.
- Chao, Yuen-Ren. 1934. The non-uniqueness of phonemic solutions of phonetic systems. *Bulletin of the Institute of History and Philology* 4:363–397.
- Charles-Luce, Jan. 1985. Word-final devoicing in German and the effects of phonetic and sentential contexts. *Journal of the Acoustical Society of America* 77.S1:S85–S85.
- Cheng, Chin-Chuan. 1968. *Mandarin Phonology*. Urbana: University of Illinois at Urbana-Champaign dissertation. (Published as *A Synchronic Phonology of Mandarin Chinese*. The Hague: Mouton, 1973.)
- Cheng, Chin-Chuan. 1973. *A Synchronic Phonology of Mandarin Chinese*. The Hague: Mouton.
- Cheon, Sang Yee & Victoria B. Anderson. 2008. Acoustic and perceptual similarities between English and Korean sibilants: implications for second language acquisition. *Korean Linguistics* 14: 41–64.
- Chiang, Wen-yu. 1992. *The Prosodic Morphology and Phonology of Affixation in Taiwanese and Other Chinese Languages*. Newark: University of Delaware dissertation.
- Chomsky, Noam & Morris Halle. 1968. *The Sound Pattern of English*. Cambridge: MIT Press.
- Clements, George Nick. 2003. Feature economy in sound systems. *Phonology* 20:287–333.
- Duanmu, San. 2007. *The Phonology of Standard Chinese*. Oxford: Oxford University Press.
- Dupoux, Emmanuel, Kazuhiko Kakehi, Yuki Hirose, Christophe Pallier & Jacques Mehler. 1999. Epenthetic vowels in Japanese: a perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance* 25:1568–1578.
- Goto, Hiromu. 1971. Auditory perception by normal Japanese adults of the sounds ‘L’ and ‘R’. *Neuropsychologia* 9:317–323.
- Gussenhoven, Carlos. 1999. Illustrations of the IPA: Dutch. *Handbook of the International Phonetic Association*, 74–77. Cambridge: Cambridge University Press.
- Hall, Kathleen Currie. 2009. *A Probabilistic Model of Phonological Relationships from Contrast to Allophony*. Columbus: Ohio State University dissertation.
- 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* 29:303–327.
- 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.
- Hockett, Charles F. 1942. A system of descriptive phonology. *Language* 18:3–21.
- Inkelas, Sharon. 1995. The consequences of optimization for underspecification. *Proceedings of the Annual Meeting of the North East Linguistic Society* 25:287–302. Amherst: GLSA, Department of Linguistics, University of Massachusetts.

- Itô, Junko & Armin Mester. 2001. Covert generalizations in Optimality Theory: the role of stratal faithfulness constraints. *Studies in Phonetics, Phonology and Morphology* 7:273–299.
- Iverson, Gregory K. & Ahnong Lee. 2006. Perception of contrast in Korean loanword adaptation. *Korean Linguistics* 13:49–87.
- 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* 38:127–136.
- Jongman, Allard, Ratree Wayland & Serena Wong. 2000. Acoustic characteristics of English fricatives. *Journal of the Acoustical Society of America* 108:1252–1263.
- Kager, René. 1999. *Optimality Theory*. Cambridge: Cambridge University Press.
- Kim, Hyunsoon. 2009. Korean adaptation of English affricates and fricatives in a feature-driven model of loanword adaptation. *Loan Phonology*, ed. by Andrea Calabrese & W. Leo Wetzels, 155–180. Amsterdam & Philadelphia: John Benjamins.
- Ladefoged, Peter & Ian Maddieson. 1996. *The Sounds of the World's Languages*. Oxford & Malden: Blackwell.
- Li, Fangfang, Jan Edwards & Mary E. Beckman. 2007. Spectral measures for sibilant fricatives of English, Japanese, and Mandarin Chinese. *Proceedings of the 16th International Congress of Phonetic Sciences (ICPS 16)*, ed. by J. Trouvain and W. J. Barry, 917–920. Saarbrücken: Universität des Saarlandes.
- Lin, Shou. 1982. V in Mandarin. *Han Yu Xue Xi* 6:1–5.
- Lin, Yen-hwei. 1989. *Autosegmental Treatment of Segmental Processes in Chinese Phonology*. Austin: University of Texas at Austin dissertation.
- Lin, Yen-Hwei. 2007. *The Sounds of Chinese*. Cambridge & New York: Cambridge University Press.
- Lisker, Leigh. 2001. Hearing the Polish sibilants [s š ś]: phonetic and auditory judgements. *To Honour Eli Fischer-Jørgensen (Travaux du Cercle Linguistique de Copenhague XXXI)*, ed. by Nina Grønnum & Jørgen Rischel, 226–238. Copenhagen: C.A. Reitzel.
- Lu, Yu-An. 2014. Mandarin fricatives redux: the psychological reality of phonological representations. *Journal of East Asian Linguistics* 23:43–69.
- McCarthy, John J. 2005. Taking a free ride in morphophonemic learning. *Catalan Journal of Linguistics* 4:19–55.
- McGuire, Grant Leese. 2007. *Phonetic Category Learning*. Columbus: Ohio State University dissertation.
- 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.
- Pisoni, David B. 1973. Auditory and phonetic memory codes in the discrimination of consonants and vowels. *Perception & Psychophysics* 13.2:253–260.
- Prince, Alan & Paul Smolensky. 2004. *Optimality Theory: Constraint Interaction in Generative Grammar*. Malden: Blackwell.
- 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: Cambridge University Press.
- Steriade, Donca. 2007. Contrast. *The Cambridge Handbook of Phonology*, ed. by Paul de Lacy, 139–157. Cambridge & New York: Cambridge University Press.

- Suh, Yunju. 2009. Perception of English voiceless alveolar and postalveolar fricative before /i/ by Korean speakers. Poster presented at the 2nd Acoustical Society of America (ASA) Special Workshop on Speech: Cross-language Speech Perception and Variations in Linguistic Experience, May 22–23, 2009. Portland, Oregon, USA.
- Swadesh, Morris. 1934. The phonemic principle. *Language* 10:117–129.
- Trager, George L. 1934. The phonemes of Russian. *Language* 10:334–344.
- Trubetzkoy, Nikolai Sergeevich. 1969. *Principles of Phonology*, translated by Christiane A. M. Baltaxe. Berkeley & Los Angeles: University of California Press.
- Wan, I-Ping. 2010. Phonological experiments in the study of palatals in Mandarin. *Journal of Chinese Linguistics* 38:157–174.
- Werker, Janet F., John H. V. Gilbert, Keith Humphrey & Richard C. Tees. 1981. Developmental aspects of cross-language speech perception. *Child Development* 52.1:349–355.
- Werker, Janet F. & Chris E. Lalonde. 1988. Cross-language speech perception: initial capabilities and developmental change. *Developmental Psychology* 24:672–683.
- Werker, Janet F. & John S. Logan. 1985. Cross-language evidence for three factors in speech perception. *Perception & Psychophysics* 37.1:35–44.
- Wu, Yuwen. 1994. *Mandarin Segmental Phonology*. Toronto: University of Toronto dissertation.
- Yip, Moira. 1996. Lexical optimization in languages without alternations. *Current Trends in Phonology: Models and Methods*, ed. by Jacques Durand & Bernard Laks, 757–788. CNRS, Paris-X and University of Salford: University of Salford Publications.

[Received 8 May 2014; revised 21 October 2014; accepted 11 January 2015]

Department of Foreign Languages and Literature
National Chiao Tung University
F319, Humanities Bldg. 2, 1001 University Road
Hsinchu 30010, Taiwan
yuanlu@nctu.edu.tw

語音分布與構詞互換於音韻關係之影響

盧郁安

國立交通大學

本文探討音韻分布與構詞互換現象對於語言使用者定義音位的影響，研究方法採用前人所建立測試語言感知及語言處理之實驗方式—語音相似度評比，研究主題包含舌尖及舌面摩擦音 [s] 與 [ʃ] 在三個不同語言裡面的分布及表現：（一）英語中的 [s] 與 [ʃ] 能區辨意義、無互補分布現象也無法從環境中預測，（二）韓語中的 [s] 與 [ʃ] 不僅呈現互補分布，也存在有構辭上的互換現象，（三）中文裡的 [s] 與 [ʃ] 呈現互補分布但無構詞互換現象。相似度評比實驗的結果顯示，英語及中文使用者對於舌尖及舌面摩擦音聽覺感知相異度較韓語使用者為大，此實驗結果支持兩音如有互換現象，即能加強語言使用者視兩音為同一音位。

關鍵詞：音韻關係，音韻分布，構詞互換，舌面摩擦音，相似度評比