

Mandarin fricatives redux: the psychological reality of phonological representations

Yu-An Lu

Received: 27 September 2011 / Accepted: 14 March 2013 / Published online: 10 September 2013
© Springer Science+Business Media Dordrecht 2013

Abstract This study contributes to the long-standing debate on the phonological representation of Mandarin palatals. The controversy results from the fact that the palatals [tɕ, tɕ^h, ɕ] are restricted to specific contexts and do not occur in the same contexts as three other sets of Mandarin consonants: the velars [k, k^h, x], the dentals [ts, ts^h, s], and the retroflexes [ʈʂ, ʈʂ^h, ʂ]. The debate has focused on the question of whether, in pursuit of an economical phoneme inventory, palatals should be derived from some other set of underlying sounds and, if so, with which series the palatal sounds should be identified. This paper reports on the results of experimental investigation of the perception and processing of the Mandarin fricatives [ɕ] and [s]. These two sounds are in complementary distribution in Mandarin and have been considered by various researchers to be allophonic variants of the same phoneme category. The results of two tasks, similarity ratings and discrimination of sounds on a continuum, suggest that even though the distribution of [s] and [ɕ] is predictable in Mandarin, Mandarin speakers do not necessarily treat the two sounds as variants of the same phoneme category.

Keywords Mandarin palatals · Economy · Fricatives · Phonological representation · Contrast · Allophony

1 Introduction

This study takes on the long-standing debate on the phonological representation of Mandarin palatals, which was first discussed in Yuen-Ren Chao's famous article in

Y.-A. Lu (✉)
National Chiao Tung University, F319 Humanities Bldg. 2, 1001 Ta-Hsueh Road, Hsinchu 30010,
Taiwan
e-mail: yuanlu@nctu.edu.tw

1934, *On the Non-uniqueness of Phonemic Solutions of Phonetic Systems*. Due to the fact that the palatals [tɕ, tɕ^h, ɕ] do not occur in the same contexts as three other sets of Mandarin consonants—the velars [k, k^h, x], the dentals [ts, ts^h, s], and the retroflexes [tʂ, tʂ^h, ʂ]—the debate has focused on the question of whether palatals should be derived from some other set of underlying sounds and, if so, with which series the palatal sounds should be identified.¹ This study employed two tasks, similarity ratings and discrimination of sounds on a continuum, to test the perception and processing of the Mandarin fricatives [ɕ] and [s]. These two sounds are in complementary distribution in Mandarin and have been considered by various researchers to be allophonic variants of the same phoneme category. The results suggest that, in spite of the predictable distribution of [s] and [ɕ], Mandarin speakers do not necessarily analyze the two sounds as variants of the same phoneme category. The following section addresses the bases of the controversy and lays out the previous analyses of Mandarin palatals.

2 Previous analyses of Mandarin palatals

Mandarin Chinese has three palatal sounds [tɕ, tɕ^h, ɕ] that are in complementary distribution with the velars [k, k^h, x], the dentals [ts, ts^h, s], and the retroflexes [tʂ, tʂ^h, ʂ]. The palatals occur before the high front vowels [i, y] and glides [j, ɥ] while the other series (dentals, velars, and retroflexes) occur elsewhere (before [u, ə, a, w]).

(1) Complementary distribution of Mandarin fricatives (Duanmu 2007, p. 31)

tɕ	tɕ ^h	ɕ	before high-front vowels [i/y] or glides [j/ɥ] (e.g., [ɕi] ‘wash’; [ɕja] ‘blind’; [ɕjo] ‘rest’; [ɕje] ‘crab’)
k	k ^h	x	
ts	ts ^h	s	never before [i/y] or [j/ɥ] (e.g., [sa] ‘spread’; [so] ‘gather’)
tʂ	tʂ ^h	ʂ	

The traditional/classic definition of contrast has relied heavily on the distribution of sounds. If no minimal pairs involving two given sounds can be found because the sounds never occur in the same context, the sounds may be considered variants of the same phoneme category. Sounds that are in complementary distribution are most often considered members of a single phoneme category (e.g., Swadesh 1934; Hockett 1942; Bloch 1948; Trubetzkoy 1969). Using this criterion, palatals might reasonably be analyzed as allophonic variants of one of the other series since they are not found in the same contexts as either dentals, velars, or retroflexes. Furthermore, if one assumes (following both structuralist (e.g., Hockett 1942) and

¹ The sounds [tɕ, tɕ^h, ɕ] are referred to as ‘palatals’ in this paper following the literature on the phonological status of these sounds (e.g., Hartman 1944; Hockett 1947; Chao 1968; Cheng 1973; Duanmu 2007). However, note that these sounds are described phonetically as palatalized post-alveolar or as alveolo-palatal (Ladefoged and Maddieson 1996, p. 150).

traditional generative approaches (Chomsky and Halle 1968)) that the inventory of phonemes should be dictated by the principle of *economy*, the pursuit of the minimal set of phonemes/features provides “pressure to eliminate the palatals as phonemes, and derive them from one of the other series” (Yip 1996, p. 770).

Despite the restricted distribution of palatals, not all researchers have agreed that the palatals should be identified with another consonant series. Cheng (1973), for instance, concludes that palatals should be considered distinct 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, p. 40). Yip (1996) came to the same conclusion—that the palatals should be distinct from the other series in underlying representations—based on the Optimality Theory notion of lexicon optimization, which posits that “learners will naturally internalize the forms closest to the surface, absent paradigm pressure [systematic morphological alternation] to do otherwise” (Yip 1996, p. 757).

Furthermore, even for those researchers who agree that surface palatals should be derived from some other series, it is not clear which series the palatals should be identified with. The following analyses of the palatals have been proposed in the literature:

(2) Analyses of Mandarin palatal sounds

a. Surface palatals derived from underlying velars or dentals

/k, k^h, x/ → [tɕ, tɕ^h, ɕ] e.g., Cheng (1968)
 /ts, ts^h, s/

b. Surface palatals derived from underlying velars

/k, k^h, x/ → [tɕ, tɕ^h, ɕ] e.g., Chao (1934), Xue (1986),
 Lin (1989), Chiang (1992), Wu (1994)

c. Surface palatals derived from underlying dentals

/ts, ts^h, s/ → [tɕ, tɕ^h, ɕ] e.g., Hartman (1944), Duanmu (2007)

d. Surface palatals derived from underlying palatals

/tɕ, tɕ^h, ɕ/ → [tɕ, tɕ^h, ɕ] e.g., Tung (1954), Cheng (1973), Yip (1996)

The claim summarized in (2a), that some palatals should be derived from underlying dentals and others from underlying velars (Cheng 1968), is based in etymological relationships. Palatals arose by means of two historical processes, *velar palatalization* and *dental sibilant palatalization* (Dong 1958; Cheng 1973)—thus, “Historically, some palatals come from the dental sibilants, others from the velar series” (Cheng 1973, p. 37).²

² The analysis of [tɕ, tɕ^h, ɕ] as underlying /tɕ, tɕ^h, ɕ/ has not been proposed in the literature presumably because the two series are etymologically unrelated (Chao 1934).

Noting that language learners do not have access to diachronic data, Chao (1934), along with other researchers (Lin 1989; Chiang 1992; Wu 1994), argues that [te, te^h, e] should be identified with the velars /k, k^h, x/. The arguments come from data from two sources, word games and onomatopoeia, both of which show palatal–velar alternations. Chao (1931) reports on a word game in which the sequence [ai.k] is infixed inside a syllable between the onset and the rhyme (e.g., [ma]→[mai.ka]), as shown in (3a–c). However, when the vowel of the original syllable is high, the infixed consonant is [tɛ] rather than [k], as in (3d).

(3) [k]~[tɛ] alternation (Chao 1931, 1934)

- a. ma → mai.ka
 b. t^ha → t^hai.ka
 c. k^hoŋ → k^hwai.koŋ
 d. liŋ → ljai.tɛiŋ

Similarly, the onomatopoeic expressions illustrated in (4) 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 (4a), but in (4b), [ɛ] appears before the front vowel and [x] before the back vowel.

(4) Onomatopoeia CV → Ci li Cu lu (Chao 1934, 1968)

- a. t^hi li t^hu lu ‘slurping’
 b. ɛi li xu lu ‘eating fast’

However, Cheng (1973) argues that these alternating patterns may be a historical residue which does not reflect synchronic phoneme categorization.

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 [ɛ] is actually a surface realization of the consonant–glide combination (CG) /sj/ (/sa/→[sa] vs. /sja/→[ɛa]; /so/→[so] vs. /sjo/→[ɛo]). Duanmu summarizes every possible CG combination, as in (5), in which a minus sign indicates a missing CG.

(5) Possible CG combinations (Duanmu 2007, p. 28)

	[j]	[w]	[ɥ]
Labial	+	–	–
Dental	+	+	+
Velar	–	+	–
Retroflex	–	+	–

He argues that the gaps occur when an articulator in the feature geometry would be involved in both members of the sequence, as in (6).

(6) Articulator analysis of CG combinations (Duanmu 2007, p. 28)

	[j] Dorsal	[w] Labial	[ɥ] Dorsal-Labial
Labial	+	–	–
Dental	+	+	+
Velar	–	+	–

As a result, sequences of velars and high-front glides (e.g., */kj/ and */kɥ/) are impossible combinations due to a principle that Duanmu calls articulator dissimilation, as in (7).

(7) Articulator dissimilation (Duanmu 2007, p. 32): Identical articulators cannot occur in succession.

Along the same lines, Duanmu attributes the missing retroflex-[j] and retroflex-[ɥ] combinations to articulatory factors as well: “In a retroflex the tongue tip is curled back, which tends to push the tongue body back, yet [j] and [ɥ] require the tongue body to be fronted” (Duanmu 2007, p. 30). In other words, the velars and the retroflexes are not compatible with high front glides.

Duanmu further argues that the realization of /sj/ as [ɕ] occurs because there is only one slot in the onset, which C and G must share.³ He also strengthens his argument that [ɕ] should be derived from /sj/ by noting a variety of Mandarin Chinese in which the CG combination is pronounced as [sʝ] instead of [ɕ]. From the distribution of glides and the correspondence between [sʝ] and [ɕ] across dialects, Duanmu concludes that the palatals should be identified with the dentals. However, one can argue that these articulatory principles against some CG combinations (i.e., velar-glide and retroflex-glide combinations) only reflect the restrictions on the surface forms and do not necessarily lead to the conclusion that the palatals are derived from underlying dentals.

The preceding discussion concerned arguments based solely on language-internal evidence. To resolve the phonological grouping of these sounds in complementary distribution, Wan (2010) attempted to investigate Mandarin speakers’ psychological analysis of the palatal sounds using four types of experimental probes to determine which series—dentals, velars, or retroflexes—the participants identified most closely with the palatals. In the first experiment, participants heard three sequences of stimuli, each of which contained an onset (a dental, retroflex, or velar), a rhyme (e.g., [ɥən]), and a full syllable (e.g., [tɕ^hɥən]), as exemplified in (8).

(8) Onset similarity experiment (Wan 2010)

Onset	Rhyme	Syllable
(a) [tɕ ^h] 3000 ms ISI	[ɥən] 1000 ms ISI	[tɕ ^h ɥən]
(b) [tɕ ^h] 3000 ms ISI	[ɥən] 1000 ms ISI	[tɕ ^h ɥən]
(c) [k ^h] 3000 ms ISI	[ɥən] 1000 ms ISI	[tɕ ^h ɥən]

³ The glides in other CG combinations are analyzed as secondary articulations, C^j, C^w, C^ɥ, that share the same onset slot.

Participants were asked to pick the most natural sequence from the three sequences. Note that the rhymes in each sequence contained a high front glide [ɥ], and the onset of each full syllable was a palatal [tɕ^h]; however, the palatal onset [tɕ^h] was absent from the single consonants played to the participants. Wan argued that if participants favored the sequences that included, before the syllable with a palatal onset, a certain type of single consonant (dental, retroflex, or velar) over the other combinations, this would indicate that that series is more closely related to the palatals and thus more likely to share the same underlying representation (Wan 2010). The results showed that the participants favored the sequences containing dental single consonants (i.e., choice (a)) significantly more often than the sequences with retroflexes or 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, because these experiments employed tasks that directly compared the similarity among the palatals, dentals, velars, and retroflexes, one can argue that the results only show that the palatals are perceptually more similar to the dentals than to the other series but do not establish that palatals should be derived from underlying dentals. In other words, the tasks might have simply reflected greater intrinsic acoustic similarity between the palatals and dentals as opposed to the other series rather than the phonological status of the palatals in the internalized grammar of the participants.

The above attempts to analyze Mandarin palatals as deriving from one of the other series of sounds in complementary distribution are based largely on the assumption that sounds with predictable distribution should be assigned to a single phoneme category (Hockett 1942; Chomsky and Halle 1968; Clements 2003). Under this assumption, the grouping of the palatals with another series in complementary distribution is inevitable. However, in phonological theories such as Optimality Theory (OT), on which Yip's (1996) analysis of Mandarin palatals is based, economy is generally assumed to play a much more limited role. Most researchers in OT assume no restrictions on the content of underlying representations, stated as *Richness of the Base*:

- (9) *Richness of the Base*: no constraints hold at the level of underlying forms (Kager 1999, p. 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. Furthermore, the principle that underlying representations should be as close as possible to surface representations (Stampe 1972; Prince and Smolensky 1993) favors the mapping of surface palatals to underlying palatals unless evidence from morphological alternation states otherwise (Inkelas 1995; Yip 1996). Mandarin palatals lack exactly this kind of evidence from alternations due to Mandarin's lack of affixation.

This paper thus strives to investigate whether predictable distribution, such as the complementary distribution of Mandarin palatals, forces sounds to map onto the same underlying representation. In other words, is the assumption of economy in the phoneme inventory a principle that guides learners' analyses of the sound system of their native language?

To investigate this question, the Mandarin palatal [ç] and dental [s], two sounds that are claimed to be underlyingly related according to Duanmu (2007) and Wan (2010), are taken as a test case. The research question is as follows: Does the contextual predictability of [s] and [ç] force Mandarin learners to conceptualize the two sounds as variants of the same phoneme category?

To test Mandarin speakers' perception and processing of [s] and [ç], two previously established methods of testing speakers' perception and processing of sounds were employed: discrimination on a continuum and similarity rating. In the discrimination experiment, Mandarin speakers' ability to distinguish [s] and [ç] was compared to their ability to distinguish the two sounds [s] and [f], which are clearly contrastive phonemes in Mandarin. The results showed that Mandarin speakers treat [s] and [ç] similarly to [s] and [f], sounds that are clearly assigned to separate phoneme categories. In the similarity rating experiment, speakers of Mandarin were asked to rate the similarity of [s] and [ç]. Their ratings were compared to those of native speakers of Korean, in which the two sounds are not only in complementary distribution but also participate in productive morphological alternations. The results showed that Mandarin speakers treated [s] and [ç] as more different than Korean speakers did. The results of the two experiments taken together suggest that Mandarin palatals, though in complementary distribution with the other series, need not map onto the same underlying representation as one of these series, consistent with Cheng (1973) and Yip's (1996) claim.

Sections 3 and 4 below present the methods and results of the discrimination and similarity rating experiments. Section 5 considers the implications of the findings.

3 Experiment I: Discrimination on a continuum

Studies of discrimination on a continuum have shown that speakers discriminate sounds that are in contrast in their native language more successfully than sounds that are allophonic variants of a single phoneme (e.g., Lisker and Abramson 1970; Lasky et al. 1975; MacKain et al. 1981; Best et al. 1988; Werker and Lalonde 1988; Lisker 2001). For example, Werker and Lalonde (1988) investigated Hindi and English speakers' ability to discriminate place of articulation in stop consonants. Hindi contrasts three places of articulation of stops—labial, alveolar, and retroflex—while English contrasts only two within this range (labial and alveolar). To determine whether speakers' discrimination was affected by the contrasts of their native language, Werker and Lalonde synthesized an 8-step continuum from [ba] to [ɖa], manipulating formant height cues signaling place of articulation (voiced labial stop to voiced retroflex). Two groups of participants, native speakers of English and native speakers of Hindi, heard pairs of sounds that were two steps apart on the continuum, presented in an ABX paradigm. The Hindi speakers' discrimination on the continuum

from [ba] to [dʌ] showed two points at which Hindi speakers were most successful in discriminating sound pairs, corresponding to the boundaries between the three categories (labial, alveolar, retroflex). The Hindi speakers were more successful in discriminating sound pairs that fell across these boundaries than pairs within boundaries. In contrast, the English speakers' discrimination on the same continuum showed only one accuracy peak, corresponding to the boundary between the English two-way contrast in place of articulation (labial, alveolar) on the [ba] to [dʌ] continuum.

In addition to accuracy of discrimination, previous research also shows that response time is a useful measure. Response time increases as a positive function of uncertainty (Studdert-Kennedy et al. 1963; Pisoni and Tash 1974): the more uncertain the listeners are, the longer they take to respond. In a discrimination task, we expect to see shorter response times when the two sounds fall across a category boundary and longer response times when the two sounds fall within a category, where the difference between the two sounds is not contrastive and the difference is presumably less salient for the listeners.

The discrimination experiment was designed to investigate Mandarin speakers' ability to discriminate sounds along a continuum from [s] to [ɛ]. If predictability of distribution forces learners to map sounds in complementary distribution onto the same underlying category, we expect that Mandarin speakers analyze [s] and [ɛ] as belonging to a single category. Along these lines, in the discrimination results, we expect to see low accuracy throughout on the continuum and in general equal response times since the difference between the sounds on the continuum would presumably pose equal difficulty for the listeners. On the other hand, if predictability of distribution does not force sounds in complementary distribution to be mapped onto the same underlying category, we expect to see evidence for a category boundary on the continuum in the form of improved discrimination of sounds from different sides of the boundary. At the same time, we expect to see shorter response times for pairs of sounds lying on different sides of the category boundary.

3.1 Methodology

Two eight-step continua were synthesized to test Mandarin listeners' discrimination, one from [s] to [ɛ] and the other from [f] to [s]. In Mandarin, [f] and [s] can occur in the same environment, and the substitution of one sound for the other may signal lexical differences (e.g., minimal pair /fǎ/ 'hair' vs. /sǎ/ 'spill'). For this continuum, we therefore expect one point at which the Mandarin speakers are most successful in discriminating sound pairs. This will serve as a baseline for a two-category response pattern. If Mandarin speakers analyze [s] and [ɛ] as members of the same category, we expect to find no evidence of a category boundary for the [s]-[ɛ] continuum. Alternatively, if Mandarin speakers analyze [s] and [ɛ] as members of separate categories, like [s] and [f], we expect evidence of a category boundary, manifested as an accuracy peak along the continuum, similar to that found for [f]-[s].

3.1.1 Participants

20 Mandarin speakers (1 male, 19 female, aged 20–22) were recruited at National Chiao Tung University in Taiwan for course credit or payment. All participants were native speakers of Taiwanese Mandarin. On a language background questionnaire, 14 participants reported that they spoke another language as well (12 speakers of Taiwanese Southern Min and 2 speakers of Hakka; see Appendix 1 for sample questionnaire). Their average self-rating of English ability was 4.4 on a 7-point scale. None reported any hearing deficiencies.

3.1.2 Design and materials

Due to Mandarin phonotactic restrictions, [s] and [ɕ] cannot be compared in identical vowel contexts.⁴ Therefore, only the frication portion of syllables [si] and [ɕi] was used in synthesizing the continuum. The endpoints of [s] and [ɕ] were spliced out using Praat software (Boersma 2001) from [si] and [ɕi] syllables spoken by a trained female phonetician whose native language is Mandarin. 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 acoustic descriptions of the spectral properties of the selected fricatives are listed in (10).

(10) Acoustic descriptions of the endpoint stimuli:⁵

	[s]	[ɕ]
Centroid frequency	8285.73 Hz	5982.45 Hz
Standard deviation	1487.63 Hz	1080.37 Hz
Skewness	−0.96	0.23
Kurtosis	5.74	4.68

The endpoints were synthesized proportionally to create an eight-step continuum using Audacity (<http://audacity.sourceforge.net/>). Step 1 was 100% [s], step 2 was 85.7% [s] and 14.3% [ɕ], and additional steps were synthesized as in (11), following the methodology in Suh (2009).⁶

⁴ It has been shown that the difference between the productions of Mandarin [s] and [ɕ] can be described in terms of two dimensions (Li 2008)—their centroid frequency and the onset F2 frequency (second formant frequency taken at the onset of the following vowel). [s] has a higher centroid frequency (around 8,000–9,000 Hz) than [ɕ] (around 4,600–7,800 Hz) while [ɕ] in general exhibits higher F2 frequency values than [s] (Li et al. 2007).

⁵ The spectral properties of the two selected fricatives fall within the range of the properties of Mandarin [s] and [ɕ] described in Li (2008), except for the standard deviation measurement.

⁶ The synthesizing process in this study did not manipulate along the acoustic dimensions mentioned in fn. 4. Instead, all the acoustic properties of [s] and [ɕ] were retained. The intervals between the two endpoints were created by overlapping different proportions of the endpoint sound tracks.

(11) Eight-step continuum from [s] to [ɛ]

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

The length of the stimuli was 270 ms. The average intensity of the stimuli was scaled to 56 dB SPL (Sound Pressure Level), the averaged intensity of the endpoints [s] and [ɛ], using Praat software. The [f]-[s] continuum was synthesized in the same way.

The experiment employed an ABX discrimination paradigm.⁷ Twelve two-step pairs (6 pairs from each continuum; steps 1–3, steps 2–4, steps 3–5, etc.) were presented randomly for each participant, using E-Prime software (v2.0; Psychological Software Tools, Pittsburgh, PA), with the members of each pair presented in each of four orders (ABB, ABA, BAA, BAB). Listeners heard each of the ABX trials (6 pairs × 4 orders × 2 continua = 48) twice in each of the 2 blocks (48 × 2 repetitions × 2 blocks = 192). The experimental variables are shown in (12).

(12) Discrimination experimental design

	Continua	[s]-[ɛ] continuum [f]-[s] continuum
Within-subject variable	Pairs (every two-step apart pairing)	s - - - - - ɛ steps 1–3, 2–4, 3–5, 4–6, 5–7, 6–8 f - - - - - s
Dependent variable	Accuracy Response time	0 = incorrect, 1 = correct in milliseconds

3.1.3 Procedure

The participants took part in the experiment individually using a computer that was connected to a keyboard with keys labeled “1” and “2.”⁸ All stimuli were presented binaurally over headphones at a comfortable listening level. An inter-stimulus interval

⁷ The editors pointed out that since the stimuli involved only the frications in isolation from any surrounding vowels, one might wonder if the subjects were responding through their grammar of Mandarin or alternatively via some more general sound perception mechanism. To facilitate a phonological level of processing in compensation of the lack of linguistic environments in the stimuli, an ABX paradigm was used, instead of an AX paradigm, to increase the memory load and to avoid an acoustic level of processing (McGuire 2009).

⁸ 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.

(ISI) of 500 ms was used. Participants were presented with written instructions in Mandarin on the computer screen saying that they would hear three sounds per trial and be asked to judge whether the third sound was the same as the first sound or the second sound. There were two blocks for the experiment with a break between the blocks. Participants had 4,000 ms to respond before the next trial started. The participants completed a 10-trial practice (randomly chosen from the test stimuli) and had the opportunity to ask questions before proceeding to the experiment. The experiment lasted approximately 10 min.

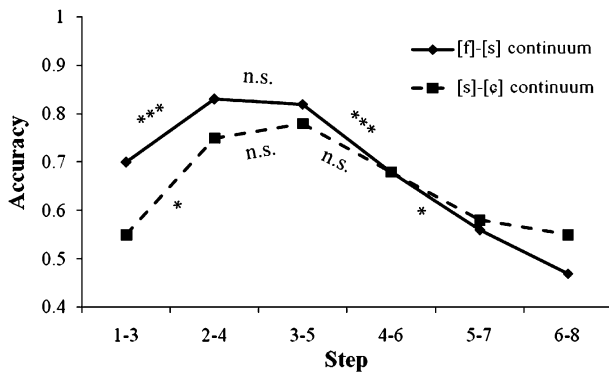
3.2 Results

The accuracy of the two continua with standard deviation in parentheses is shown in (13) and illustrated in (14).

(13) Discrimination accuracy

Step	1–3	2–4	3–5	4–6	5–7	6–8
[f]-[s] continuum	.70 (.138)	.83 (.121)	.82 (.137)	.68 (.183)	.56 (.140)	.47 (.127)
[s]-[ɕ] continuum	.55 (.146)	.75 (.16)	.78 (.17)	.68 (.16)	.58 (.13)	.55 (.14)

(14) Accuracy for [f]-[s] and [s]-[ɕ] continua



The x axis of (14) identifies the fricative pairs that were presented, and the y axis represents the accuracy with which each pair was discriminated. As we can see, the results provided evidence of a boundary for both continua, located somewhere between steps 2–5 for the [f]-[s] continuum, indicated by the solid line, and between steps 2–6 for the [s]-[ɕ] continuum, indicated by the dashed line. A repeated-measures analysis of variance (ANOVA) confirmed this observation.⁹ For the [f]-[s] continuum, there was a main effect of Pair ($F(5,95) = 22.149, p < .001$), which was indicative of the difference in accuracy across sound pairs two steps apart. Pairwise comparisons showed that steps 2–4 vs. 3–5 were not significantly different ($p = .859$), but steps 1–3

⁹ The assumption of normality was met for running ANOVAs. See Appendix 2.

vs. 2–4, and steps 3–5 vs. 4–6 were significantly different (both $p < .001$). This suggests that the perceptual boundary falls between steps 2–5 on the [f]-[s] continuum, as shown in (14). (*: $p < .05$; **: $p < .01$; ***: $p < .001$; *n.s.*: not significant).

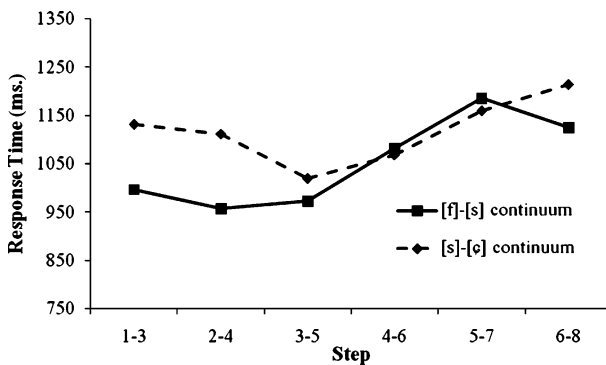
Similarly, another repeated-measures ANOVA was run on the [s]-[ε] continuum. There was also a main effect of Pair ($F(5,95) = 9.610$, $p < .001$). Pairwise comparisons among steps 2–4, 3–5, and 4–6 were not significant (all $p > .05$). On the other hand, steps 1–2 vs. 2–4, and 4–6 vs. 5–7 were significantly different (both $p < .05$), as shown in (14).¹⁰

The results from the response times are also consistent with a perceptual boundary on the two continua. The response times for the two continua, with standard deviations in parentheses, are shown in (15) and illustrated in (16). The x axis of (16) indicates the positions of the fricative pairs on the continuum, and the y axis represents the response times in milliseconds.

(15) Discrimination response times

Step	1–3	2–4	3–5	4–6	5–7	6–8
[f]-[s] continuum	996 (155)	956 (152)	972 (215)	1081 (287)	1185 (326)	1125 (275)
[s]-[ε] continuum	1131 (289)	1111 (319)	1019 (204)	1068 (249)	1159 (268)	1214 (390)

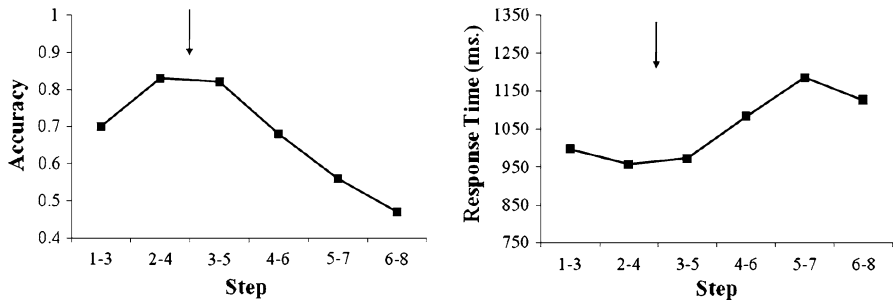
(16) Response times for [f]-[s] and [s]-[ε] continua



¹⁰ An anonymous reviewer pointed out that the boundary in both cases ([s]-[ε] and [f]-[s]) is skewed towards the left end of the continuum. One would expect the boundary to be in the middle given that the stimuli pairs were presented randomly and the steps on the continuum were of equal acoustic distance. Although this study cannot account for the location of the boundary, it has been shown that human perception is non-linear (Johnson 1997). For example, a change in an acoustic manipulation is not equivalent to a similar change perceptually. Although the steps on the continua used in this study were manipulated proportionally with equal distance from the end points, the participants might be more sensitive to the acoustic change on the left end of the continua. The important point here is that a perceptual boundary is present somewhere along *both* continua.

For the [f]-[s] continuum, indicated by the solid line, we see shorter response times in the beginning of the continuum and longer response times towards the end of the continuum. Furthermore, the pairs for which the participants took less time to respond to corresponded to the pairs that the participants were more successful in discriminating. The corresponding pattern is shown clearly in (17) when we put the accuracy and the response time results side by side.

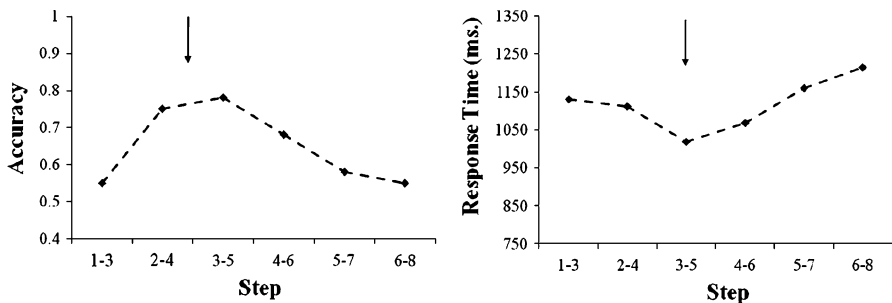
(17) [f]-[s] continuum accuracy and response time results



The valley of the response time results in the right panel, indicated by the arrow, corresponded nicely to the peak of accuracy in the left panel. In other words, the participants took less time responding to the pairs that they perceived more accurately while they took longer to respond to the pairs with lower accuracy.

Crucially, we observe a similar response time pattern on the [s]-[ç] continuum, as shown in (18).

(18) [s]-[ç] continuum accuracy and response time results



Just like the results of the [f]-[s] continuum, the valley of the response time results of the [s]-[ç] continuum in the right panel, indicated by the arrow, corresponded to the peak of the accuracy on the left panel. The results for the Mandarin participants parallel the findings reported in other studies in which response time serves as a positive function of uncertainty (Pisoni and Tash 1974): when the two sounds crossed a category boundary, the response times were shorter; when the two sounds compared fell within a category, the response times were longer.

Taken together, the accuracy and response time results suggest that the discrimination of [s] and [ɕ] was not different from the discrimination of two uncontroversially contrastive sounds (i.e., [s] and [f]) by the Mandarin listeners. This finding suggests that the complementary distribution of [s] and [ɕ] does not necessarily force native speakers to map the two sounds onto the same phoneme category.

3.3 Discussion

A possible complicating factor in this experiment is the fact that more than half of the participants (12 out of 20; cf. Sect. 3.1.1) reported that they were also speakers of Taiwanese Southern Min. As a reviewer pointed out, the correspondences between palatal and dental sibilants between Taiwanese Southern Min and Mandarin are not always consistent, as shown in (19).

(19) Mandarin palatal and dental sibilants in Taiwanese Southern Min

		Mandarin	Taiwanese Southern Min	Gloss
[s]	→	[sɿ]	[ɕi]	‘four’
[ɕ]	→	[ɕ ¹ ao]	[ɕ ¹ ɿ]	‘little’
[s]	→	[san]	[sã]	‘three’
[ɕ]	→	[ɕi]	[se]	‘west’

Mandarin and Taiwanese Southern Min bilinguals might be aware of the fact that some occurrences of [s] in Mandarin correspond to [ɕ] in Taiwanese, and some occurrences of [ɕ] correspond to [s]. Although the two sounds are also in complementary distribution in Taiwanese and have been analyzed as variants of the same phoneme (Chung 1996, p. 14), the different correspondences in Mandarin and Taiwanese in terms of the palatal and dental sibilants might enable bilingual speakers to be aware of the [s] and [ɕ] sound difference. However, a post-hoc repeated-measures ANOVA on the accuracy results including language background as a variable (Language Background [Mandarin, Mandarin-Taiwanese bilinguals] × Pair) did not show an effect ([s]-[ɕ] continuum: $F(1, 16) = .008$, $p = .930$); [f]-[s] continuum: $F(1, 16) = 3.225$, $p = .091$). This suggests that the discrimination patterns of the participants with Taiwanese Southern Min background were not different from the patterns of the participants without such background. The statistical results with language background as a variable are summarized in (20).¹¹

¹¹ The language background of Hakka was not taken into account since the number of Hakka speakers (2 out of 20) was not sufficient to run a statistical analysis.

(20) Steps \times Language Background

[f]-[s] continuum	[s]-[ɕ] continuum
***Step	***Step
$F(5, 80) = 36.602, p < .001$	$F(5, 80) = 15.624, p < .001$
Background	Background
$F(1, 16) = 3.225, p = .091$	$F(1, 16) = .008, p = .930$
Step \times Background	Step \times Background
$F(5, 80) = 1.843, p = .114$	$F(5, 80) = .572, p = .721$

To summarize, this experiment tested the ability of Mandarin speakers to discriminate pairs of sounds from an eight-step continuum from [s] and [ɕ] and another continuum from [f] and [s] as a comparison. The experiment was designed to compare their discrimination of [s]-[ɕ] with discrimination of clearly contrastive sounds, [f]-[s]. The accuracy and response time results suggested that Mandarin speakers perceived both the [s]-[ɕ] continuum and the [f]-[s] continuum in terms of two categories, consistent with the view that the complementary distribution of [s] and [ɕ] does not force the two sounds to map onto the same phoneme category.

4 Experiment II: Similarity rating

In similarity rating tasks, listeners have exhibited a tendency to rate sounds that represent allophonic variants of a single phoneme category as more similar than sounds representing separate phoneme categories (Harnsberger 2001; Boomershine et al. 2008; Babel and Johnson 2010; Johnson and Babel 2010). For example, Boomershine et al. (2008) tested native English and Spanish speakers' similarity judgments of [ð], [d], and [r] using an AX paradigm. [ð] and [d] are contrastive in English (e.g., *they* [ðeɪ] vs. *day* [deɪ]) but allophonic in Spanish, due to a process whereby intervocalic voiced stops are spirantized (e.g., [d]onde 'where' but *de* [ð]onde 'from where'). In contrast, [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]) become taps intervocalically preceding an unstressed syllable (e.g., *ride* [raɪd], but *rider* [raɪrə]). In Boomershine et al.'s study, participants were asked to rate the similarity of two sounds taken from the VCV sequences [ada], [ara], [aða], [idi], [iri], [iði], [udu], [uru], 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 show a clear native language effect, with English speakers rating [d] and [r] as most similar, but Spanish speakers rating [ð] and [d] as most similar, reflecting the phonological relationships of the three sounds in their native language.

The second set of experiments in this study used a similarity rating task to compare the similarity ratings for [s] and [ɕ] by native speakers of Mandarin and native speakers of Korean (cf. Wan 2010). Korean was chosen as a point of comparison because the facts of Korean provide strong support for analyzing these two sounds as members of a single phoneme category. First, as in Mandarin, the two

sounds are in complementary distribution, with [ɛ] occurring only before the high front vowel/glide [i/j], and [s] occurring elsewhere (Sohn 1999; Iverson and Lee 2006; Kim 2009), illustrated in (21).

(21) Complementary distribution of Korean [s] and [ɛ]

a.	[ɛi]	‘poem’
b.	[ɛikan]	‘time’
c.	[ɛjamp ^h u]	‘shampoo’
d.	[ɛjap]	‘shop’
e.	[ɛjup ^h Δ]	‘super’
f.	[ɛjo]	‘show’
g.	[sal]	‘flesh’
h.	[sul]	‘alcohol’
i.	[se]	‘bird’

Furthermore, many morphemes exhibit alternation between the two sounds arising when affixation places a final [s] before a high front vowel, as shown in (22).¹²

(22) Morphological alternation of [s] and [ɛ] in Korean

a. /nas/	[nas-e]	‘sickle-locative’
	[naɛ-i]	‘sickle-nominative’
b. /kos/	[kos-e]	‘place-locative’
	[koɛ-i]	‘place-nominative’
c. /pus/	[pus-e]	‘writing brush-locative’
	[puɛ-i]	‘writing brush-nominative’

4.1 Methodology

In this experiment, native speakers of Mandarin and Korean were asked to rate the target sounds [s] and [ɛ] in terms of similarity. Since [s] and [ɛ] are in complementary distribution and participate in regular and productive morphological alternations in Korean, we therefore expect Korean listeners to rate [s] and [ɛ] as very similar, due to the status of these sounds as allophonic variants in Korean. The goal of this experiment is to see how Mandarin listeners rate the similarity between the two target sounds. If the Mandarin speakers’ ratings are comparable to those of Korean listeners, then [s] and [ɛ] can be considered to be allophonic variants of a single phoneme, just as in Korean. This would suggest that predictable distribution forces the two sounds to map onto the same phoneme category. If the Mandarin speakers’ ratings are different from those of the Korean listeners, this would suggest that the predictable distribution of [s] and [ɛ] in Mandarin does not force learners to map the two sounds onto the same category while the combination of distributional predictability and morphophonological alternations in Korean does encourage learners to assign the two sounds to a single category.

¹² Note that the OT approach discussed in Sect. 2 considers that only evidence of morphological alternation can force learners to posit a single underlying representation that differs from the phonetic representation.

4.1.1 Participants

20 Mandarin (4 male, 16 female, aged 20–22) and 20 Korean (6 male, 14 female, aged 18–38) speakers participated in this experiment. Participants in the Mandarin group were recruited at National Chiao Tung University in Taiwan for course credit or payment. They were all native speakers of Taiwanese Mandarin. On a language background questionnaire, 16 participants reported that they spoke another language as well (13 speakers of Taiwanese Southern Min, and 3 speakers of Hakka). Their average self-rating of English ability was 4.6 on a 7-point scale. Participants in the Korean group were recruited at SUNY Stony Brook University and received payment for their participation. They all received up to a high school education in South Korea before coming to SUNY Stony Brook University for undergraduate or graduate education. Their average self-rating of English ability was 4.65 on a 7-point scale. None reported any hearing deficiencies.

4.1.2 Design and materials

Twelve disyllabic VCV stimuli were used in this set of experiments, composed of the target fricatives [s, ɕ] along with two other fricatives [f, h] as controls, embedded in three vowel contexts [a_a], [i_i], and [u_u] (4 fricatives × 3 vowel contexts = 12 VCV stimuli). Note that the sound /f/ does not exist in Korean, and some of the stimuli contained illicit sequences according to the phonotactics of Mandarin and Korean: *[si], *[ca], and *[eu]. The tokens were produced by a trained male phonetician whose native language is Mandarin because he was able to produce the tested fricatives natively 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. I selected one instance of each VCV so that the tokens were approximately matched on pitch and duration. (23) shows the pitch of the vowels (V1 mean: 116.17 Hz, standard deviation: 2.48 Hz; V2 mean: 116.25 Hz, standard deviation: 2.09 Hz), and (24) shows the vowel and fricative durations of the stimuli (total duration mean: 729.67 ms, standard deviation: 34.57 ms).

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

	V1	V2
aea	118	117
afa	120	118
aha	119	117
asa	114	116
iei	113	115
ifi	114	114
ihi	113	114
isi	114	112
ueu	118	118
ufu	118	119
uhu	116	118
usu	117	117

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

	V1	Fric	V2	Total
aca	225	198	330	753
afa	301	142	333	775
aha	277	128	313	717
asa	266	162	337	765
iei	255	201	320	776
ifi	262	152	329	743
ihi	278	137	305	720
isi	243	182	309	734
ueu	194	213	315	722
ufu	226	169	291	685
uhu	231	152	288	671
usu	203	196	297	695

The average intensity for each token was scaled to 65 dB SPL using Praat software, the rough average of the intensity of all the tokens.

The design followed closely that of Boomershine et al. (2008), as shown in (25).

- (25) Similarity rating design

Between-subject variable	Language	→Korean, Mandarin
Within-subject variable	Fricative Pair	→[s-ɕ], [s-f], [s-h] [e-f], [e-h] [f-h]
Dependent variable	Rating score	→1(similar) to 5(different)

This experiment employed an AX paradigm comparing pairs of fricatives in three vowel contexts (6 fricative pairs \times 2 orders \times 3 vowel contexts = 36). The listeners heard each of the AX trials three times in the 3 blocks (36 \times 3 blocks = 108) with an ISI of 1,000 ms between A and X. Participants had a maximum of 5,000 ms to respond before the next trial started.

4.1.3 Procedure

Participants took part in the experiment individually, using a computer that was connected to a keyboard with 5 keys labeled from 1 to 5. The participants were presented with written instructions in their native language on the computer screen saying that they would hear a pair of sounds over headphones and be asked to rate how similar those sounds were on a scale of 1–5, where 1 was ‘very similar’ and 5 was ‘very different.’ The pairs were presented in different random orders for each participant, using E-Prime software. The participants completed a 9-trial practice, randomly chosen from the test stimuli, and had the opportunity to ask questions before proceeding to the experiment. The experiment lasted approximately 20 min.

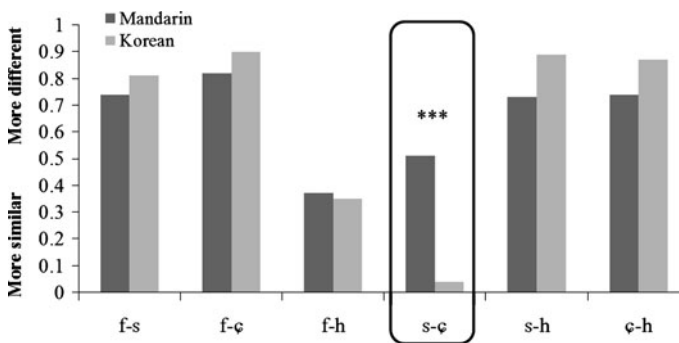
4.2 Results

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’ (see Appendix 2 for the distribution of the transformed results). The normalized results are shown in (26) and (27). In (27), the x axis represents the different fricative pairs, and the y axis represents the normalized z-scores.

(26) Similarity rating normalized results—means

Pair \ Language	f-s	f-ç	f-h	s-ç	s-h	ç-h
Mandarin	0.74	0.82	0.37	0.51	0.73	0.74
Korean	0.81	0.90	0.35	0.04	0.89	0.87

(27) Similarity rating normalized results—figure



We can see from (27) that the ratings were very similar for the two languages, with the exception of the target pair [s-ç], indicated by the solid box. A repeated-measure ANOVA (Language [Mandarin, Korean] \times Pair [f-s, f-ç, f-h, s-ç, s-h, ç-h]) was performed to interpret the results. The analysis showed that there was a main effect of Pair ($F(5,38) = 73.545, p < .001$). In other words, the ratings differed for different fricative pairs. There was also a significant Pair by Language interaction ($F(5,190) = 15.077, p < .001$), meaning that the ratings for pairs of fricatives were statistically different depending on the native language of the participants. An effect of Language in the [s-ç] pair was also found ($F(1,38) = 36.692, p < .001$), meaning that the ratings of the [s-ç] pair from the Mandarin group were statistically higher than those from the Korean group. The fact that the Korean speakers rated [s-ç] as more similar than did the Mandarin speakers suggests that the Korean speakers were more likely to analyze these two sounds as variants of the same category than were the Mandarin speakers. These results are consistent with the results of the previous experiments suggesting that

the predictable distribution of the Mandarin [s] and [ɕ] does not force Mandarin speakers to treat the two sounds as variants of the same category.

4.3 Discussion

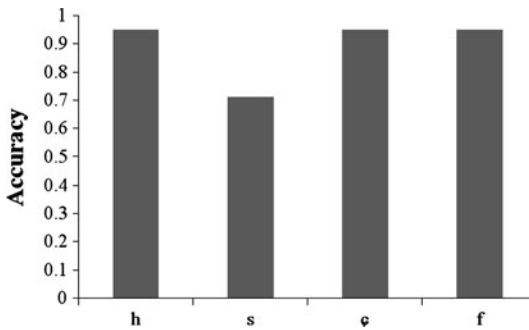
In this section I explore a possible alternative explanation of the greater differences in the Mandarin vs. Korean similarity ratings of [s] and [ɕ]. An anonymous reviewer pointed out that the stimuli in this experiment included illegal sequences in both Mandarin and Korean (e.g., [ueu] and [æa]; cf. section 4.1.2). It is possible that Mandarin listeners rated the illegal sequences, [ueu]/[æa], as relatively more different from [usu]/[asa] than did Korean listeners because the post-alveolar [ɕ] in these illegal contexts might have been misperceived as the other Mandarin post-alveolar fricative, retroflex [ʂ]. This misperception would create *legal* sequences in these vowel contexts (i.e., [u_u] and [a_a]), and these sequences would clearly contrast with [s] in Mandarin (e.g., [su] ‘crispy’ vs. [ʂu] ‘lose’; [sa] ‘spread’ vs. [ʂa] ‘sand’). Such misperception, however, is not possible for the Korean speakers since the other coronal fricatives, tense [sʰ] and [ɕʰ], are subject to the same phonotactic restrictions as [s] and [ɕ].

To rule out this possibility, a followup identification experiment was conducted to verify what Mandarin listeners identify [ɕ] as in these illegal contexts. Another 10 Mandarin speakers (2 male, 8 female, aged 23–34) participated voluntarily in the follow-up identification experiment. They were all native speakers of Taiwanese Mandarin who had received up to a college education in Taiwan before coming to the United States. Their average self-rating of English ability was 5 on a 7-point scale. None reported any hearing deficiencies. The same stimuli were used as in the similarity rating experiment. The participants heard each of the stimuli four times (12 VCV stimuli \times 4 = 48) through headphones at a comfortable listening level and in a different random order for each participant, using E-Prime software. The participants were presented with written instructions on the computer screen asking them to indicate whether they heard a [s], [ɕ], [ʂ], [h], or [f] by pressing the keys on a keyboard with keys labeled “ㄌ”, “ㄒ”, “ㄑ”, “ㄒ”, “ㄒ”, in Zhuyin Fuhao/Bopomofo.¹³ The participants completed a 6-trial practice run randomly chosen from the stimuli and had the opportunity to ask questions before proceeding to the experiment. The experiment lasted approximately 3 min.

The crucial question is whether the Mandarin listeners misperceived [ɕ] in illegal vowel contexts ([a_a] and [u_u]) as the retroflex [ʂ], which could subsequently cause them to rate the [s-ɕ] pair as more different than did the Korean listeners. The results are shown in (28). The *x* axis indicates the fricatives being identified, and the *y* axis indicates the accuracy of identification. As we can see, except for [s], there is a ceiling effect in the responses, suggesting that the participants were very successful in identifying the other three fricatives.

¹³ Zhuyin Fuhao/Bopomofo is a phonetic system taught to school-age children before the standard non-alphabetic writing system in Taiwan.

(28) Identification results



The results showed that the Mandarin speakers were very successful in identifying [ɕ] tokens as palatal (0.95 accuracy). Only 6 out of 120 instance of [ɕ] were identified as [s] or [ʂ]. Furthermore, while the identification of [s] was less accurate, [s] was classified as [ʂ] in only 10 cases. The other 25 cases of misidentification involved classifying [s] as [ɕ], and these cases were all embedded in the [i_i] vowel context. The misperception of [s] as [ɕ] in the [i_i] context does not seem to be surprising since the vowel [i] provides a pre-palatal context. While Mandarin speakers might reasonably have been biased by the phonotactic restriction in their native language to misperceive [s] as [ɕ] in this context, this followup identification experiment suggests that the possibility of the Mandarin listeners misperceiving [ɕ] as [s] in these illegal contexts was not likely and that the similarity rating results did suggest a different phonological grouping of [s] and [ɕ] in Mandarin and Korean.

To summarize this section, the results of the experiment investigating how listeners of Mandarin and Korean rated the similarity of [s] and [ɕ] showed that Mandarin listeners rated [s-ɕ] as significantly more different from each other than did Korean listeners. These results are consistent with the results of the discrimination experiment, which suggested that Mandarin listeners need not map the two sounds in complementary distribution, [s] and [ɕ], as members of the same phoneme category.

5 Conclusion

The studies in this paper were designed to investigate how Mandarin native speakers conceptualize the relationship between the dental and palatal fricatives, two sounds in complementary distribution. Would predictable distribution, in the absence of allomorphic alternation, force Mandarin speakers to analyze the sounds as member of the same category?

The experiments conducted in this paper, as opposed to the ones in Wan (2010), avoided direct comparisons among the series in complementary distribution. Evidence from a discrimination experiment revealed that Mandarin listeners perceived a continuum from [s] to [ɛ] similar to a [s]-[f] continuum. Mandarin speakers also judged these sounds as less similar than did Korean speakers, for whom the two sounds are clearly members of the same category. Taken together, the categorical perception on the [s]-[ɛ] continuum and the phonemic-like judgment on the similarity rating task suggest that predictable distribution does not force Mandarin speakers to analyze [s] and [ɛ] as variants of the same category.

The results add to the ongoing debate concerning the status of these sounds in Mandarin and also shed light on the definition of phonological relationships. The controversy surrounding the analysis of Mandarin palatals results from the fact that the three palatals [tɕ, tɕ^h, ɕ] do not occur in the same context as the three other series: the dentals [ts, ts^h, s], the velars [k, k^h, x], and the retroflexes [ʈʂ, ʈʂ^h, ʂ]. The challenge is that distributional predictability is often taken as the diagnostic for phoneme assignment and that “in an abstract analysis, economy of phoneme inventory supplies pressure to eliminate the palatals as phonemes, and derive them from one of the other series” (Yip 1996, p. 770). The results from this study suggest that sounds in complementary distribution, like [s] and [ɛ] in Mandarin, need not map onto the same underlying representation. These results pose challenges for phonological theories that rely heavily on distribution in defining phonological relationships and in which the concept of economy is taken to be essential in phoneme analysis, such as the traditional structuralist approach (e.g., Swadesh 1934; Hockett 1942; Bloch 1948; Trubetzkoy 1969) and the traditional generative approach (Chomsky and Halle 1968; Kenstowicz and Kisseberth 1979; Clements 2003). In contrast, these results are more consistent with theories in which economy in underlying phoneme inventories is not a driving factor, such as Optimality Theory (Prince and Smolensky 1993), in which the underlying representations are not restricted (cf. *Richness of the Base* (9)), and the learner is not necessarily assumed to remove all predictable information from underlying representations.

Furthermore, the contrast between the similarity ratings of the Mandarin speakers vs. the Korean speakers suggests that, apart from the traditional definition of contrast/allophony, based in predictability of distribution, the lack of *morphological alternation* may also dispose learners to assign sounds that never alternate to different categories. Mandarin [s] and [ɛ], though in complementary distribution, do not alternate due to Mandarin’s lack of affixation and stringent restrictions on possible syllable structures. Korean, on the other hand, shares with Mandarin the predictable distribution of [s] and [ɛ] but differs from Mandarin in that these sounds also participate in regular and productive morphological alternation (cf. (21) and (22)). The results of the similarity rating experiment, in which Mandarin speakers rated [s] and [ɛ] as more different than did the Korean speakers, suggest that the additional evidence from alternation that the Korean speakers are exposed to had an effect. That is to say, multiple factors (i.e., distribution and alternation) may contribute to the formation of sound categories. And if phonological relationships are built up by different criteria, then the relationship between two sounds should

not be a clear-cut one. In other words, phonological relationships should be gradient. These results 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 (Goldsmith 1995; Crowley 1998; Kristoffersen 2000; Moulton 2003; Ladd 2006; Rose and King 2007; Scobbie and Stuart-Smith 2008). In this camp, Hall (2009) 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. 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, p. 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 suggested in this paper, in which multiple factors (e.g., distribution and alternation) may interact in determining the phonological relationships among sounds of a language.

To conclude, I identify two areas for future research, one on the status of Mandarin fricatives and the other on the definition of phonological relationships. First, because this paper compared Mandarin 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–velar and palatal–retroflex) to see if the same results hold. Second, the findings in the studies suggest that different criteria may contribute to the formation of phoneme category and that phonological relationships should be gradient, and not absolute.¹⁴ It will also be left for future research to investigate the relative contributions of different factors in defining phonological relationships.

Acknowledgments I would like to thank Ellen Broselow, Kathleen Currie Hall, Marie Huffman, and the *JEAL* reviewers and editors for all their comments and ideas. I would also like to thank Yuwen Lai for her help in running the experiment. This work was supported by NSF grant BCS-07460227 to Ellen Broselow, Marie Huffman and Nancy Squires, and the Chiang Ching-Kuo Foundation Dissertation Fellowship to the author.

¹⁴ These factors include the ones examined here (distribution and alternation) and phonetic similarity, lexical distinction, and orthography. The last factor is also relevant in the analysis of Mandarin palatals as the editors pointed out that the Zhuyin Fuhao/Bopomofo writing system, a phonetic system used before learning the ideographic writing system, utilizes different letters for each of the sounds in complementary distribution. To determine whether the orthography affected the results presented here, one area for future study is to see if the same results hold for pre-school-age children.

Appendix 1: Example questionnaire

Participant number: _____

Email: _____

Age: _____

Gender: _____

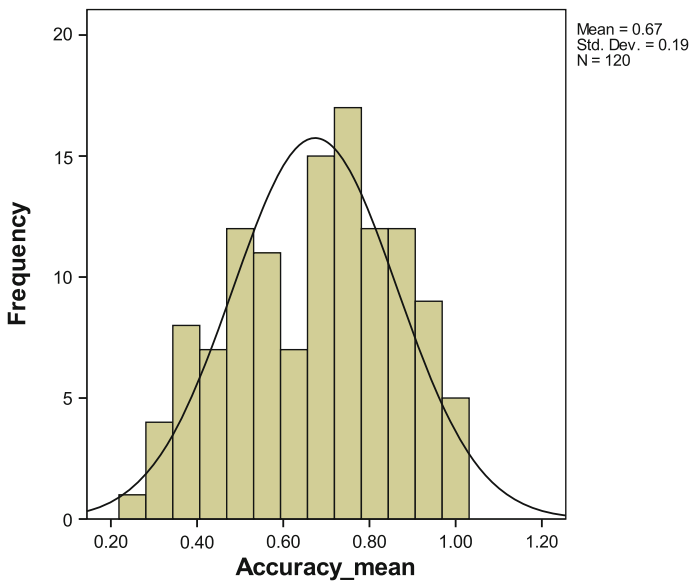
Language Group:

- What languages do you speak?
- Self-rated English ability

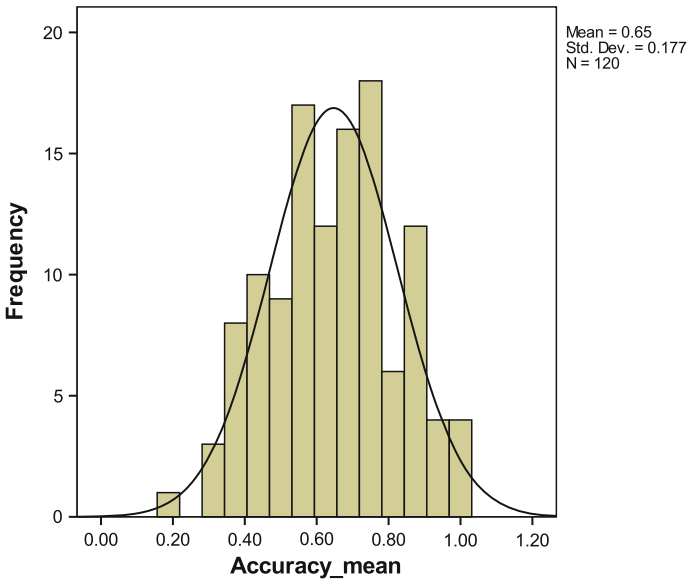
	Very bad					Very good	
	1	2	3	4	5	6	7
Listening	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Speaking	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Reading	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Writing	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Overall	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Appendix 2

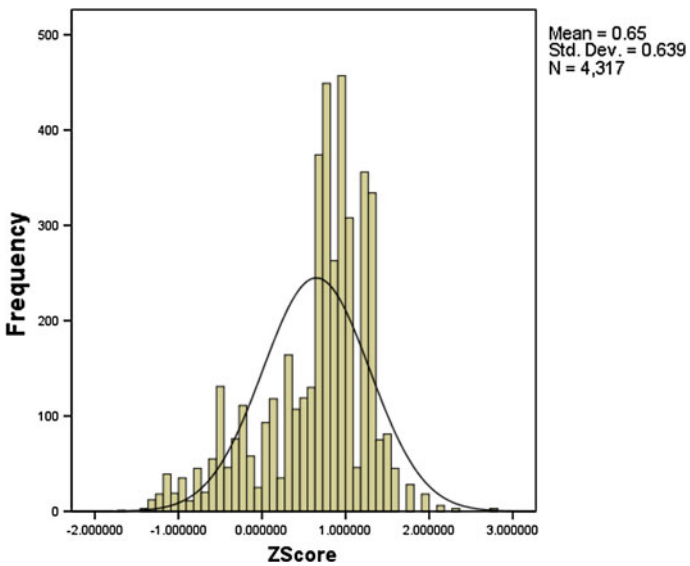
Distribution of the [f]-[s] continuum accuracy results



Distribution of the [s]-[ɛ] continuum accuracy results



Distribution of the similarity rating z-scores



References

- Babel, Molly, and Keith Johnson. 2010. Accessing psycho-acoustic perception with speech sounds. *Laboratory Phonology* 1: 179–205.
- Best, Catherine T., Gerald W. McRoberts, and 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: 345–360.
- Bloch, Bernard. 1948. A set of postulates for phonemic analysis. *Language* 24: 3–46.
- Boersma, Paul. 2001. Praat, a system for doing phonetics by computer. *Glott International* 5(9/10): 341–345.
- Boomershine, Amanda, Kathleen Currie Hall, Elizabeth Hume, and Keith Johnson. 2008. The impact of allophony versus contrast on speech perception. In *Contrast in phonology*, ed. Peter Avery, Elan Dresher, and Keren Rice, 143–172. Berlin: 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.
- Chao, Yuen-Ren. 1968. *A grammar of spoken Chinese*. Berkeley and Los Angeles: University of California Press.
- Cheng, Chin-Chuan. 1968. *Mandarin phonology*. Ph.D. dissertation, University of Illinois at Urbana-Champaign.
- Cheng, Chin-Chuan. 1973. *A synchronic phonology of Mandarin Chinese*. Berlin: De Gruyter Mouton.
- Chiang, Wen-Yu. 1992. *Prosodic phonology and morphology of affixation in Chinese*. Ph.D. dissertation, University of Delaware.
- Chomsky, Noam, and Morris Halle. 1968. *The sound pattern of English*. Cambridge, MA and London: The MIT Press.
- Chung, Raung-fu. 1996. *The segmental phonology of Southern Min in Taiwan*. Taipei: The Crane Publishing Co., Ltd.
- Clements, G.N. 2003. Feature economy in sound systems. *Phonology* 20: 287–333.
- Crowley, Terry. 1998. The voiceless fricatives [s] and [h] in Erromangan: One phoneme, two, or one and a bit? *Australian Journal of Linguistics* 18: 149–168.
- Dong, Shaowen. 1958. *Yuyin changtan [Introduction to phonetics]*. Beijing: Wenhua Jiaoyu Chubanshe.
- Duanmu, San. 2007. *The phonology of standard Chinese*. New York: Oxford University Press.
- Goldsmith, John. 1995. Phonological theory. In *The handbook of phonological theory*, ed. John Goldsmith, 1–23. Cambridge, MA: Blackwell.
- Hall, Kathleen Currie. 2009. A probabilistic model of phonological relationships from contrast to allophony. Ph.D. dissertation, 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: 303–327.
- 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.
- Hockett, Charles F. 1947. Peiping phonology. *Journal of the American Oriental Society* 67: 253–267.
- Inkelas, Sharon. 1995. The consequences of optimization for underspecification. In *Proceedings of the 25th meeting of North East Linguistic Society*, ed. E. Buckley, and S. Iatridou, 287–302. Amherst: GLSA.
- Iverson, Gregory K., and Ahnong Lee. 2006. Perception of contrast in Korean loanword adaptation. *Korean Linguistics* 13: 49–87.
- Johnson, Keith. 1997. *Acoustic and auditory phonetics*. West Sussex, UK: Blackwell.
- Johnson, Keith, and 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.
- Kager, René. 1999. *Optimality theory*. Cambridge: Cambridge University Press.
- Kenstowicz, Michael, and Charles Kisseberth. 1979. *Generative phonology: Description and theory*. New York: Academic Press.
- Kim, Hyunsoon. 2009. Korean adaptation of English affricates and fricatives in a feature-driven model of loanword adaptation. In *Loan phonology*, ed. Andrea Calabrese and W. Leo Wetzels, 155–180. Amsterdam: John Benjamins.

- Kristoffersen, Gjert. 2000. *The phonology of Norwegian*. Oxford: Oxford University.
- Ladd, D. Robert. 2006. "Distinctive phones" in surface representation. In *Laboratory phonology 8*, ed. Louis M. Goldstein, D.H. Whalen, and Catherin T. Best, 3–26. Berlin: Mouton de Gruyter.
- Ladefoged, Peter, and Ian Maddieson. 1996. *The sounds of the world's languages*. Cambridge: Blackwell.
- Lasky, Robert E., Ann Syrdal-Lasky, and 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: 215–225.
- Li, Fangfang. 2008. The phonetic development of voiceless sibilant fricatives in English, Japanese and Mandarin Chinese. Ph.D. dissertation, The Ohio State University.
- Li, Fangfang, Jan Edwards, and Mary Beckman. 2007. Spectral measures for sibilant fricatives of English, Japanese, and Mandarin Chinese. In *Proceedings of the XVth international congress of phonetic sciences*, ed. J. Trouvain, and W.J. Barry, 917–920. Dudweiler: Pirrot.
- Lin, Yenhwei. 1989. Autosegmental treatment of segmental processes in Chinese phonology. Ph.D. dissertation, University of Texas at Austin.
- Lisker, Leigh. 2001. Hearing the Polish sibilants [s̥ š ś]: Phonetic and auditory judgements. In *travaux du Cercle Linguistique de Copenhague XXXI. To honour Eli Fischer-Jørgensen*, ed. Nina Grønnum and Jørgen Rischel, 226–238. Copenhagen: C.A. Reitzel.
- Lisker, Leigh, and Arthur S. Abramson. 1970. The voicing dimensions: Some experiments in comparative phonetics. In *Proceedings of the sixth international congress of phonetic sciences*, ed. B. Hála, M. Romportl, and P. Janota, 563–567. Prague: Academia.
- MacKain, Kristine S., Catherine T. Best, and Winifred Strange. 1981. Categorical perception of English /t/ and /l/ by Japanese bilinguals. *Applied Psycholinguistics* 2: 369–390.
- McGuire, Grant. 2009. A brief primer on experimental design for speech perception. Ms. Department of Linguistics, University of California, Santa Cruz.
- Moulton, Keir. 2003. Deep allophones in the Old English laryngeal system. *Toronto Working Papers in Linguistics* 20: 157–173.
- Pisoni, David B., and Jeffrey Tash. 1974. Reaction times to comparisons within and across phonetic categories. *Perception and Psychophysics* 15: 285–290.
- Prince, Alan, and Paul Smolensky. 1993. *Optimality theory: Constraint interaction in generative grammar*. Malden: Blackwell.
- Rose, Sharon, and Lisa King. 2007. Speech error elicitation and co-occurrence restrictions in two Ethiopian semitic languages. *Language and Speech* 50: 451–504.
- Scobbie, James M., and Jane Stuart-Smith. 2008. Quasi-phonemic contrast and the indeterminacy of the segmental inventory: Examples from Scottish English. In *Contrast in phonology: Theory, perception and acquisition*, ed. Peter Avery, B.Elan Dresher, and Keren Rice, 87–114. Berlin: Mouton.
- Sohn, Ho-Min. 1999. *The Korean language*. Cambridge: Cambridge University Press.
- Stampe, David. 1972. How I spent my summer vacation [A dissertation in natural phonology]. Ph.D. dissertation, University of Chicago.
- Studdert-Kennedy, Michael, Alvin M. Liberman, and Kenneth N. Stevens. 1963. Reaction time to synthetic stop consonants and vowels at phoneme centers and at phoneme boundaries. *The Journal of the Acoustical Society of America* 35: 1900–1900.
- Suh, Yunju. 2009. Perception of English voiceless alveolar and postalveolar fricative before /i/ by Korean speakers. Paper presented at Acoustical Society of America (ASA), 2nd ASA special workshop on speech, Portland, OR.
- Swadesh, Morris. 1934. The phonemic principle. *Language* 10: 117–129.
- Trubetzkoy, Nikolai Sergeevich. 1969. *Principles of phonology* (trans. Baltaxe, Christiane A.M.). Berkeley and Los Angeles: University of California Press.
- Tung, Tung-ho. 1954. *Zhongguo Yuyin Shi [A phonological history of Chinese]*. Taipei: Chonghua Wenhua Chubanshe.
- Wan, I-Ping. 2010. Phonological experiments in the study of palatals in Mandarin. *Journal of Chinese Linguistics* 38: 157–174.
- Werker, Janet F., and Chris E. Lalonde. 1988. Cross-language speech perception: Initial capabilities and developmental change. *Developmental Psychology* 24: 672–683.
- Wu, Yuwen. 1994. *Mandarin segmental phonology*. Ph.D. dissertation, University of Toronto.
- Xue, Fengsheng. 1986. *Guo yu yin xi jie xi [An anatomy of the pekingese sound system]*. Taipei: Taiwan Xuesheng Shuju.
- Yip, Moira. 1996. Lexical optimization in languages without alternations. In *Current trends in phonology: Models and methods*, ed. J. Durand, and B. Laks, 757–788. Paris-X/Salford: CNRS/University of Salford Publications.