Tonal bilingualsim: the case of two related Chinese dialects
Wu, J.

Citation
Wu, J. (2015, July 2). *Tonal bilingualsim: the case of two related Chinese dialects*. *LOT dissertation series*. LOT, Utrecht. Retrieved from https://hdl.handle.net/1887/33727

Author: Wu, Junru
Title: Tonal bilingualism : the case of two closely related Chinese dialects
Issue Date: 2015-07-02

2 Interlingual Two-to-One Mapping of Tonal Categories

Abstract

Both Standard Chinese (SC) high- and low-rising tones sound like the rising tone in Jinan Mandarin (JM). The role of this two-to-one interlingual tonal mapping was investigated in speech production, speech perception, and lexical access. Statistical modeling suggests that both SC rising tones overlap with the JM rising tone, with the high-rising SC tone having a better interlingual category-goodness. A semantic priming experiment focused on whether and how SC rising tones activate JM words in native bilinguals' mental lexicon and to what extent the interlingual category-goodness exerts its influence on speech comprehension. The bilinguals made auditory lexical decisions from a list including JM real words, JM non-words, and SC pseudo-words which sound like other JM real words. The SC pseudo-words (primes) were followed by the JM real words (targets). These targets were semantically related or unrelated to the JM real words hinted at by the SC primes. The SC pseudo-words ended with either high-rising or low-rising tones. The results showed that both high-rising and low-rising final SC pseudo-words were accepted as JM real words with a higher-than-chance probability. The activation of JM lexical items from SC tones was asymmetrical. SC pseudo-words ending with high-rising tones were more likely and more quickly recognized as real JM words, compared with their counterparts ending with low-rising tones. However, when accepted as JM real words, the final low-rising and final high-rising pseudo-words were equivalent in their semantic priming effects. These results support overlapping divisions of the tonal acoustic space according to the language mode in lexical access and some discreteness in the stage between lexical activation and semantic activation.

2.1 Introduction

Two different phonemes in one language matching to one and the same phoneme in the other language is a common phenomenon in bilingualism. For instance, Dutch and German learners have difficulty distinguishing English /æ/ and /ɛ/ because they only have /ɛ/ whose acoustic distribution primarily overlaps with, but is still different from the English /ɛ/ (Bohn & Flege, 1990; Flege, Bohn, & Jang, 1997). Similarly, Japanese learners have difficulty distinguishing English /r/ and /l/ because they only have /ɾ/ which, while also apico-alveolar, is instead tapped (Best & Strange, 1992; Miyawaki et al., 1975). Such phenomena have been extensively investigated in second-language phoneme perception, and the related confusions in lexical access have also been studied.

In lexical decision, the minimal pairs, which are not contrastive in the native language become 'pseudo-homophones' (e.g. English locket vs. rocket for Japanese listeners) and prime each other like repetitions for the same word (Anne Cutler & Otake, 2004; Dufour, Nguyen, & Frauenfelder, 2007; Pallier, Colomé, & Sebastián-Gallés, 2001). 'Near-words' constructed by replacing a phoneme (e.g. English /t/) with its confusing phoneme (e.g. English /d/) are taken as words by listeners who have difficulty distinguishing the pair in their native language (e.g. Dutch /t/ and /d/ are neutralized in word-final position[1]); for the non-native listeners in repetition priming, such ‘near-words’ embedded in a context also prime the corresponding real word (e.g. ‘half lied’ primed ‘flight’) causing ‘phantom word activation’ (Broersma & Cutler, 2008).

The two-to-one interlingual mapping can be asymmetrical. For instance, the Japanese /ɾ/ is perceptually more similar to the English /l/ than the English /r/ (Aoyama, Flege, Guion, Akahane-Yamada, & Yamada, 2004; Iverson et al., 2003). The phonetic asymmetry also affects how the phonetic representations activate lexical representations. In auditory picture-word identification, the picture of a ‘locker’ triggered more interference with the selection of the target ‘rocket’ than vice versa (A. Cutler, Weber, & Otake, 2006). A similar asymmetry was found between ‘pencil’ and ‘panda’ for Dutch-English bilinguals and Dutch learners of English (Escudero, Hayes-Harb, & Mitterer, 2008; Weber & Cutler, 2004). According to Cutler and Weber (2004, 2006), these results support two aspects of bilingual lexical access. First, the L2 acoustic input is captured by the native phonemic categories. Second, even though an inaccurate categorization may happen during the phonetic processing before lexical access, non-contrastive L2 phonemes in native perception are still stored with distinctive mental representations.

Similar two-to-one mapping patterns also abound in tonal languages and dialects. For instance, Jinan Mandarin (JM) has only one rising tone (JM T1) (Qian, 1997) but Standard Chinese(SC) has two rising tones, one high (SC T2) and one low (SC T3)(Moore & Jongman, 1997; Shen & Lin, 1991). This two-to-one tonal mapping is conditional. Specifically, this two-to-one cross-language tonal mapping is only valid in isolation and at prosodic word boundaries, where the SC high- and low-rising tones are contrastive but only one JM rising tone exists.

The two-to-one interlingual mapping of tones may nevertheless be different from the previously investigated segmental cases. First, lexical tone is a special type of phoneme in speech perception and lexical access. Although categorical perception of native tones is supported by recent neurological studies (Chandrasekaran, Krishnan, & Gandour, 2009; Xi, Zhang, Shu, Zhang, & Li, 2010), the native perception of Chinese tones in behavioral studies is ‘quasi-categorical’, neither as categorical as that of consonants nor as continuous as that of vowels (Hallé, Chang, & Best, 2004). Tonal information, compared with segmental information, is also retrieved later (Ye & Connine, 1999; Zhang & Damian, 2009; Zhang & Zhu, 2011) and involves different neuronal networks (Liang & van Heuven, 2004) in speech production. Considering lexical access, although lexical adaptation (McQueen, Cutler, & Norris, 2006; Mitterer, Chen, & Zhou, 2011) and constraining activation (Malins & Joanisse, 2010) seem to work similarly in tones and consonants, the overlap of SC tones alone induced no facilitatory priming effect in implicit priming, different from the classical effect observed by segmental primes (Chen, Chen, &

---

[1] According to (Warner, Jongman, Sereno, & Kemps, 2004), this neutralization of Dutch word-final /t/ and /d/ may be incomplete. Some sub-phonemic durational difference is maintained, which, although inconsistent in production, can still be noticed in perception.

Dell, 2002). In spite of these studies, the role of tone in perception and lexical access is little studied in bilinguals who use two tonal systems natively.

Second, aside from the particularity of tone compared with segmental phonemes, the two-to-one mapping between SC and JM rising tones seems different when regarding the overlapping of acoustic distributions. The Japanese /ɾ/ is a tap (Vance, 1987), while the English /l/ and /r/ are lateral and central approximants, respectively (Ladefoged, 2001). The north German /ɛ/ has lower F1 (less open) and higher F2 (more front) than the American English /ɛ/ (Strange, Bohn, Trent, & Nishi, 2004). Thus in both cases, the two-to-one mapping actually involves three similar but distinctive phones (i.e. the English /æ/, English /ɛ/[ɛ], and the German /ɛ/[ɛ̝]). However, is this also the case of SC-JM tonal mapping? The acoustic distribution of the JM rising tone seems to largely overlap with the acoustic distribution of both the SC high-rising and low-rising tones, although the mapping with the SC high-rising tone seems to be greater. This mapping asymmetry needs experimental verification before we make further investigations into their contribution to bilingual lexical access.

Third, different from the L2 learners in previous two-to-one mapping studies, the young SC-JM bilinguals are mostly simultaneous or early bilinguals with SC as the dominant language and received literacy education only in SC. These bilinguals thus have high proficiency in both dialects and have little difficulty in producing the correct rising tones according to the language mode and the lexical meaning.

How tonal bilinguals map tonal representations to lexical representations in speech perception and lexical access, however, needs more investigation. To investigate this two-to-one interlingual tonal mapping case, some other situations need to be controlled. On the one hand, the SC low-rising tone has different variants in non-final position. Before another low-rising tone, the SC low-rising and high-rising tones are nearly merged (Peng, 2000; Yuan & Chen, 2014) and not distinguishable in speech perception. The non-final tonal realizations both sound like the SC high-rising tone in isolation. For instance, ‘land reform’ and ‘alter’ are homophones in SC, due to the tone sandhi which happens on the first word /tʰu(low-rising) + kai(low-rising) -> tʰukai(high-rising+low-rising) = tʰu(high-rising) + kai(low-rising)/. Also, the rising part of the SC low-rising tone only appears in isolation at prosodic word boundaries (Garding, Kratochvil, Svantesson, & Zhang, 1986; Hallé et al., 2004). On the other hand, the JM rising tone realizes congruently higher in non-final positions and seems to only map to the SC high-rising tone. Thus, JM and SC rising tones have a one-to-one mapping (high-rising with high-rising) in non-final positions. To avoid the confusion from tone sandhi, SC stimuli with low-rising tone on the first syllable were avoided in the current study.

The acoustic overlapping between JM and SC rising tones gives rise to additional questions. How do the SC-JM bilinguals store the rising tones in their mental lexicon for auditory lexical access? More specifically, do they store the JM rising tone with a separate mental representation or do they store the JM rising tone integrated with the more similar SC high-rising tone? One possibility is that they store only two tonal representations, one high-rising and the other low-rising, but the categorization of the high-rising tone serves both JM and SC lexical access. If this hypothesis is true, we should expect the canonical realization of the SC low rising tone to fail matching any tonal representation in JM lexical access. The other possibility is that the bilinguals store three tonal representations. Then the theory needs to allow two tonal representations in the same mind to have overlapping acoustic distributions. This hypothesis also implies that the bilinguals can process the same tonal realization differently depending on the dialect in use. This is possible when considering the previous finding that Greek–English early sequential bilinguals gave different category-goodness ratings for the same physical stimuli depending on the language mode (Antoniou, Tyler, & Best, 2012). If this hypothesis is true, the canonical realizations of both the SC high-rising and low-rising tones could be taken as the JM rising tone in lexical access. Also, similar to the asymmetry found for both consonants and vowels (A. Cutler et al., 2006; Weber & Cutler, 2004), acoustic realizations of the SC high-rising tone, compared with that of the SC low-rising tone, should be better captured by the mental representation of the JM rising tone. Thus, with the segmental structure aligned, the SC pseudo-word with a high-rising tone should more likely be accepted as a real JM word. Also, the lexical decision of the SC high-rising pseudo-word should be faster, considering the asymmetry found in the bilingual lexical access involving asymmetrical two-to-one segmental mapping patterns (A. Cutler et al., 2006; Escudero et al., 2008; Weber & Cutler, 2004).

There is also a further question about the influence of the category-goodness. The canonical realization of the SC high-rising tone, compared with that of the SC low-rising tone, is a better exemplar of the JM rising tonal category. As discussed above, we expect this difference in category-goodness to influence the interlingual phantom activation of the JM lexical representations. However, to what extent does this category-goodness maintain its influence in speech comprehension? Researchers have a strong consensus that the phonological and semantic activation proceed in a largely parallel way in auditory speech comprehension. For instance, the ‘cohort model’ suggests that, as soon as the first phonemes of the target word are identified, the corresponding semantic information is activated to some extent (Grosjean, 1980; William Marslen-Wilson, 1984; W. D. Marslen-Wilson & Welsh, 1978). This parallel or cascading view of speech comprehension has been supported by both behavioral and neurophysiological evidence. For instance, shadowing can be performed before the utterance ends (W. Marslen-Wilson, 1973). Partial primes are enough to cause associative priming in sentential context (Zwitserlood, 1989). The semantic effect on the N400 component onsets before the offset of the eliciting word (Holcomb & Neville, 1991) and the latency of the N400 component is affected by whether the pseudo-word is different from the real word at the beginning (O'Rourke & Holcomb, 2002; Van Petten, Coulson, Rubin, Plante, & Parks, 1999). Comparing the timing of the N200 ERP components related to the monitoring of phonological and conceptual features in a Go/noGo paradigm, the former component precedes but largely overlaps with the later component (Rodriguez-Fornells, Schmitt, Kutas, & Münte, 2002). However, these findings did not distinguish lexical activation from semantic activation. If the tonal category-goodness influences lexical access, does this influence also reach the semantic level? If it does, SC pseudo-words with high-rising tone should also increase the semantic activation of the corresponding JM real word. Alternatively, if the influence of the category-goodness stops after lexical activation and does not spread to the semantic level, the SC pseudo-words with high-rising and low-rising tones should show no further difference in semantic activation, so long as they succeed in activating the JM lexical node. The strength of semantic activation can be tested with the semantic priming paradigm.

In the current study, three experiments tested the production, perception, lexical access, and semantic effects of the SC and JM rising tones. Experiment 1 compared the acoustic distribution of the two SC rising tones against the JM rising tone and verified the asymmetrical mapping in speech production. Experiment 2 tested SC tonal monolinguals' perception of the JM rising tone as SC tones and verified the asymmetrical mapping in SC speech perception. Experiment 3 tested SC-JM bilinguals' auditory lexical decision of SC pseudo-words (with high-rising and low-rising tones) as JM real words (with the JM rising tone), as well as their semantic priming effects on JM real words. The lexical decision of the SC pseudo-words as JM real words reflects the effect of interlingual tonal goodness in bilingual lexical access. We expect the SC high-rising pseudo-words to be more likely and quickly accepted as JM real words. By comparing the semantic priming effects of SC high-rising and low-rising pseudo-words we want to answer the question whether the influence of the tonal category-goodness in speech comprehension is maintained at the semantic level.

2.2 Experiment 1: acoustic distribution of JM and SC rising tones

Experiment 1 aimed at investigating the acoustic distributions of the JM rising tone and the two SC rising tones at the end of disyllabic word forms (JM words and SC pseudo-words).

2.2.1 Participants

Forty-two native JM tonal bilinguals from Jinan (16 male and 26 female, aged between 23 and 76 years, M = 40.29, SD = 17.04; seventeen SC dominant or balanced, twenty-five JM dominant) and 48 SC tonal monolinguals from Beijing (7 male and 41 female, aged between 19 and 30 years, M = 22.73, SD =2.95) participated in this experiment in exchange for payment.

2.2.2 Corpus preparation

The list of stimuli is composed of 16 final-rising disyllabic JM words and their corresponding SC pseudo-words, which either end with high-rising or low-rising SC tones (see the appendix for the full list).

The JM final-rising disyllabic words were selected from a 400-word corpus produced by 42JM speakers. The selection was controlled in that the targeted JM tonal patterns were produced by at least 88% of the JM speakers in our corpus and the JM words do not have false-friends in SC. In order to construct the corresponding SC pseudo-words, these disyllabic JM words also satisfy the criterion that their monosyllabic compositions have false-friends in SC and do not have frequent alternative pronunciation. Both JM and SC morphemes can be written with Chinese characters.

For each JM final-rising word, a pair of SC pseudo-words was constructed, which shares the same monosyllabic morpheme in the first syllable and ends with monosyllabic tonal minimal pairs carrying SC high and low-rising tones, respectively. For instance, a pair of SC disyllabic pseudo-words were constructed as ‘绳银’ /ʂəŋ in(high-rising+high-rising)/ and ‘绳引’ /ʂəŋ in(high-rising+low-rising)/,which are homophone candidates for the JM word ‘sound’ /ʂəŋ in([high-]rising+rising)/. The SC pseudo-words were then presented with Chinese characters and named by the SC tonal monolinguals.

For both the JM and SC participants, the printed stimuli were presented in a different random order for each speaker. After the speaker finished producing a word or pseudo-word, they pressed a key to see the next word. We used Praat (Boersma & Weenink, 2014) to extract pitch contours. Only pitch contours on the rhymes were extracted. A trained phonetician listened to each recording, looked at the spectrogram, and manually marked the rhyme of each syllable. Also, in this process, recordings with speech errors or recording errors were excluded from the corpus. Afterwards, the pitch contours were converted from hertz to semitones with 100Hz as the base and then transformed into z-scores based on the speakers’ means and standard deviations. This normalization removed the difference of pitch register across speakers, which is not the focus of the present study. The normalized pitch contours were then interpolated to 20 points per-syllable to remove the difference in duration.

2.2.3 Analysis and results

We built Generalized Additive Models (GAM) for the *Pitch* data, using the ‘mgcv’ package (S. Wood, 2006; S. N. Wood, 2011) in R (R_Core_Team, 2013). The data were stratified according to the JM tonal combinations (each together with their two corresponding SC tonal combinations) and fitted with separate models. In each model, a three- level factorial predictor *tone* [JM rising (JM t1), SC high-rising (SC t2), and SC low-rising (SC t3)] was included. Smooth functions were used to model non-linear functional relations between *Pitch* and the position of the point on the pitch contour (*pnt*). *Tone* (the JM or SC tone carried by the ending syllable) was included in both the fixed linear predictors and fixed smoothes. The candidates for random predictors were *itemID*, *set ID* (each JM real word and its similar SC final high-rising and low-rising pseudo-words form a set), and *Speaker*. Since *item ID* is nested under *set ID* and predictable due to the combination of *set ID* and *tone*, we built models which would otherwise be identical, including *item ID* and *set ID* together, *set ID* alone, or *item ID* alone in the random terms. The structure of the final model was decided by model comparison based on the Akaike Information Criterion (AIC) likelihood values (Sakamoto & Ishiguro, 1986). The models with the *item ID* alone turned out to be the best. Thus, *item ID* but not *set ID* was included in the random terms. After the structure of the model was decided, autocorrelation values were calculated based on the order of data points in the pitch contour and the greatest value was included as the AR1 correlation parameter to build the corresponding AR1 error model (S. Wood, 2006; S. N. Wood, 2011) but when the AR1 error model does not improve the original model, the original model was reported.

The final models for different SC tonal combinations shared the same structure. *Tone* was included as the factorial predictor; a thin-plate regression spline smooth was included to model the interaction of *pnt* and *tone*; the smoothes of the by-*Speaker* random slope of *pnt* and the smooth of by-*word ID* random slope of *pnt* were included as the random predictors.

The fitted models accounted for 76.1% of the variance in the data of the JM and SC rising + (high/low) rising tonal combination, 86.7% of the variance in the data of the JM and SC high-level + (high/low) rising tonal combination, 76.4% of the variance in the data of the JM and SC falling + (high/low) rising tonal combination. The coefficients for the parametric predictors are shown in Table 1. The number of degrees of freedom in the smooth terms and the associated F-statistics are shown in Table2. The fitted pitch contours are shown in Figure1.

As shown by the scattered contour plots in Figure 1 (van Rij, Wieling, Baayen, & van Rijn, 2015), the JM rising tone (JM T1) overlaps with both SC rising tones in the acoustic distribution, with greater overlap with the SC high-rising tone (SC T2). As shown by the estimated smoothes in Figure 1, SC final low-rising (SC T3) pseudo-words carry lower pitch contours on the second syllable than SC final high-rising (SC T2) pseudo-words and the contours of JM final-rising (JM T1) words lie in between the two SC rising tones. With the rising (JM T1 & SC T2) and high-level (JM T3 & SC T1) tones on the first syllable, although the JM tones seem to cover a slightly smaller and lower pitch register, the cross-dialect mapping is very close. However, the pitch register of JM low-falling (JM T4) on the first syllable is much lower than that of the SC falling (SC T4). Hence, it is mainly the shape of the pitch contour instead of the pitch register that is similar between the JM and SC falling + rising combinations (JM T4 + T1& SC T4 + T2/T3). Also, the shape of this JM tonal combination is more similar to the final high-rising SC counterpart than to the other tonal combinations.

Table 1. Coefficients for the linear predictors in the generalized additive model fitted to *Pitch* of JM & SC production data (** $p < 0.01$, *** $p < 0.001$).

	Predictors	Estimate	Std. Error	t value
JM T1 + T1 & SC T3 + T2/T3	(Intercept)	-0.2896	0.1545	-1.875 (n.s.)
rising + (high/low) rising:	toneSCt2	0.2642	0.2256	1.171 (n.s.)
	toneSCt3	-0.6342	0.2262	-2.804**
JM T3 + T1 & SC T1 + T2/3	(Intercept)	-0.02666	0.09758	-0.273 (n.s.)
high-level + (high/low) rising:	toneSCt2	0.41335	0.13908	2.972**
	toneSCt3	-0.15217	0.13913	-1.094 (n.s.)
JM T4 + T1 & SC T4 + T2/3	(Intercept)	-0.61862	0.08596	-7.197***
low-falling + (high/low) rising:	toneSCt2	0.51235	0.12191	4.203***
	toneSCt3	-0.09893	0.12202	-0.811(n.s.)

Table 2. Coefficients for the smooth terms in the generalized additive model fitted to *Pitch* of JM & SC production-identification data (*** $p < 0.001$).

	Smooth terms	edf	Ref.df	F
JM T1 + T1 & SC T3 + T2/T3 rising + (high/low) rising:	s(pnt):toneJMt1	16.41	16.46	680.99***
	s(pnt):toneSCt2	15.97	16.41	29.67***
	s(pnt):toneSCt3	16.70	16.83	141.81***
	s(pnt, Speaker)	708.31	763.00	353.04***
	s(pnt, itemID)	81.94	105.00	806.76***
JM T3 + T1 & SC T1 + T2/3 high-level + (high/low) rising:	s(pnt):toneJMt1	16.86	16.88	694.26***
	s(pnt):toneSCt2	16.52	16.87	50.37***
	s(pnt):toneSCt3	16.84	16.96	181.25***
	s(pnt, Speaker)	692.59	763.00	295.54***
	s(pnt, itemID)	104.15	132.00	436.13***
JM T4 + T1 & SC T4 + T2/3 low-falling + (high/low) rising:	s(pnt):toneJMt1	16.8	16.83	403.6***
	s(pnt):toneSCt2	15.2	16.32	44.64***
	s(pnt):toneSCt3	16.91	16.98	281.07***
	s(pnt,Speaker)	709.57	763.00	490.91***
	s(pnt, itemID)	153.69	186.00	67.48***

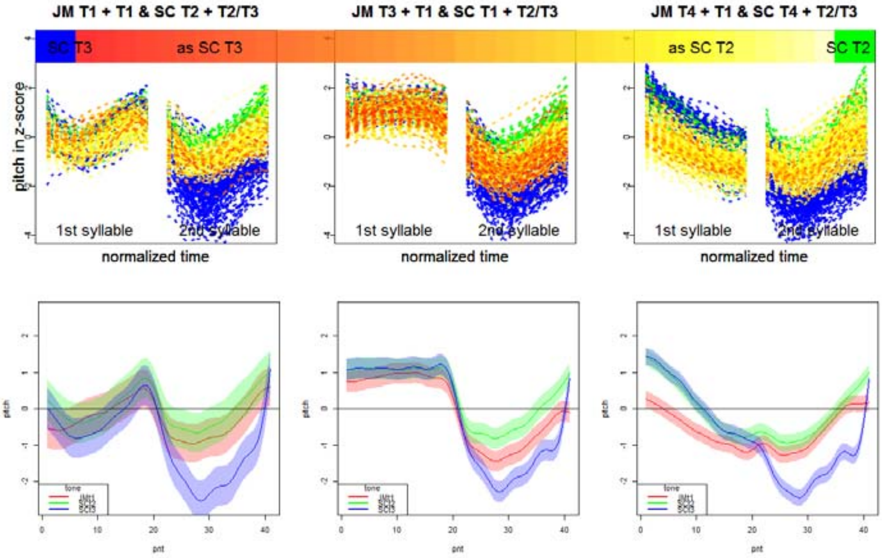

Figure 1. *Scattered contour plots (upper panel) and estimated smoothes (lower panel, random effect removed) for the JM final-rising words [JM T1, scaled from red (contour identified as SC T2) to yellow (contour identified as SC T3)], SC final-high-rising pseudo-words (SC T2, green), and S final-low-rising pseudo-words (SC T3, blue).*

2.2.4 Discussion

Experiment 1 verified that the acoustic distribution of the JM rising tone largely overlaps with both the SC high-rising and low-rising tones at word final position, with the overlap with the SC high-rising tone larger than the overlap with the SC low-rising tone. Despite the distributional overlapping, the GAM modeling showed that the three rising tones have different distributional centers, with the distributional center of the JM rising tone lying between the SC high-rising and low-rising tones.

2.3 Experiment 2: identification of JM words as SC pseudo-words

In Experiment 2, we verified the claim that the acoustic realizations of the JM rising tone can match both SC high-rising and low-rising tones in SC speech perception and investigated the relation between interlingual identification and the shape of pitch contours.

2.3.1 Participants

The forty-eight SC tonal monolinguals who participated in Experiment 1 also participated in Experiment 2.

2.3.2 Design and Stimuli

The 16 final rising disyllabic JM words (as shown in Appendix) produced by the 42 JM speakers in Experiment 1, with production errors and non-dominant variants excluded, were used as the stimuli in Experiment 2. Additionally, another SC-JM tonal bilingual who is highly proficient in both dialects produced the corresponding SC pseudo-words for each of the 16 JM words, which served as the training stimuli.

2.3.3 Procedure

The SC tonal monolingual participants performed a tonal identification task upon JM auditory stimuli using the E-Prime software (Schneider, Eschman, & Zuccolotto, 2002) in a quiet room. With the JM words presented binaurally in random order, the participants judged which of the two corresponding SC pseudo-words printed on the screen was heard. The participants had 5,000 ms to make the judgment and the following stimulus appeared 1,000 ms after the response. Thirty-two training trials with real SC pseudo-words were tested before the crucial trials with feedback. It was verified that the participants were able to identify the training SC pseudo-words with high accuracy (ACC) [grand mean ACC = 0.97 (SD = 0.17), by-item ACC ranged from 0.74 to 1, by-participant ACC ranged from 0.91 to 1].

2.3.4 Analysis

We built Generalized Additive Models (GAM) for the *Pitch* data, using the 'mgcv' package (S. Wood, 2006; S. N. Wood, 2011) in R (D. M. Bates, 2010). Separate models were built for different JM tonal combinations and smooth functions were used to model non-linear functional relations between the predictors and the measurement *Pitch*. The SC tonal monolinguals' choice for each stimuli (*Choice*), and position of the point in the pitch contour (*pnt*) were included as the fixed predictors, and *wordID*, *Speaker*, and *Participants* were included as the candidates for random predictors. The structure of the final model was decided by model comparison based on the AIC likelihood values (Sakamoto & Ishiguro, 1986). *Participant* did not reach significance or improve the model and was removed. Similar to the procedure used in Experiment 1, autocorrelation values were calculated and included as the AR1 correlation parameter to build the corresponding AR1 error model (S. Wood, 2006; S. N. Wood, 2011) but when the AR1 error model does not improve the original model, the original model was reported.

The final models for different SC tonal combinations shared the same structure. *Choice* was included as the factorial predictor; a thin-plate regression spline smooth was included to model the interaction of *pnt* and *Choice*. The smooth of the by-*Speaker* random slope of *pnt* and the smooth of the by-*word ID* random slope of *pnt* were included as the random predictors.

2.3.5 Results

The fitted models accounted for 75% of the variance in the data of the JM rising + rising tonal combination, 86.5% of the variance in the data of the JM high-level + rising tonal combination, 74% of the variance in the data of the JM low-falling + rising tonal combination. The coefficients for the parametric predictors are shown in Table 3. The numbers of degrees of freedom in the smooth terms and the associated F-statistics are shown in Table 4. The fitted pitch contours of the JM stimuli identified as final high-rising and low-rising SC pseudo-words and the corresponding difference contours are shown in Figure 2, Figure 3, and Figure 4. As shown in these figures, the differences of pitch contours between the stimuli identified as final high-rising and final low-rising (SC T2 vs. T3) SC pseudo-words are consistent. Stimuli with relatively lower pitch contours on the second syllable are more likely to be identified as carrying the SC low rising tone (SC T3) on that syllable. Also, the stimuli with higher pitch contours on the first syllable are more likely to be identified as carrying the SC low rising tone on the second syllable.

GAM models similar to those built in Experiment 1 yielded the plots in the lower panels of Figures 2, 3, and 4. As shown in these plots and the scattered contours in Figure 1, the SC final low-rising (SC T3) pseudo-words carry lower pitch contours on the second syllable than the SC final high-rising (SC T2) pseudo-words. The shape of the difference curves is consistent with the difference curves from the perceptual data of JM real words. However, the SC final low-rising pseudo-words do not always carry higher pitch contours on the first syllable for compensation. Also, it seems that the SC monolinguals' identification of JM tones as SC tones is mostly based on the shape of the pitch contour. As was also shown in the results of Experiment 1 (Figure 1), the pitch contour of JM low-falling (JM T4) is much lower than that of the SC falling (SC T4). Although the general pitch register is different, the relative pitch register of JM low-falling (JM T4), compared with that of JM rising (JM T1), makes the shape of pitch contour of the JM low-falling + rising (JM T4 + T1) combination more similar to that of the SC falling + high-rising combination. As shown in Figure 4, the difference curve between the two identified subcategories is more similar to that of the other JM tonal combinations than the corresponding SC difference curve. Also, a majority of the JM T4 + T1 stimuli were identified as SC final high-rising pseudo-words.

Table 3. Coefficients for the linear predictors in the generalized additive model fitted to *Pitch* of JM production-identification data (** $p < 0.01$, *** $p < 0.001$).

	Predictors	Estimate	Std. Error	t value
JM T1 + T1	(Intercept)	0.120992	0.288234	0.42(ns.)
rising + rising:	Choice: as T3	-0.030936	0.005596	-5.529***
JM T3 + T1	(Intercept)	0.17975	0.18029	0.997 (ns.)
high-level + rising:	Choice: as T3	-0.02345	0.00488	-4.805 ***
JM T4 + T1	(Intercept)	-0.453184	0.133084	-3.405***
low-falling + rising:	Choice: as T3	-0.041467	0.005352	-7.748***

Table 4. Coefficients for the smooth terms in the generalized additive model fitted to *Pitch* of JM production data (*** $p < 0.001$).

	Smooth terms	edf	Ref.df	F
JM T1 + T1	s(pnt):Choice: as SC T2	16.04	16.05	1082.00***
rising + rising:	s(pnt):Choice: as SC T3	16.15	16.17	269.80***
	s(pnt, Speaker)	366.20	377.00	299.30***
	s(pnt, Word ID)	30.69	35.00	829.00***
JM T3 + T1	s(pnt):Choice: as SC T2	16.61	16.62	1329.70***
high-level + rising:	s(pnt):Choice: as SC T3	16.62	16.64	551.20***
	s(pnt, Speaker)	365.82	377.00	237.80***
	s(pnt, Word ID)	38.54	44.00	640.90***
JM T4 + T1	s(pnt):Choice: as SC T2	16.62	16.64	1248.50***
low-falling + rising:	s(pnt):Choice: as SC T3	16.57	16.62	144.10***
	s(pnt, Speaker)	366.56	377.00	456.50***
	s(pnt, Word ID)	56.46	62.00	923.30***

2.3.6 Discussion

Experiment 2 verified that the JM rising tone matches both the SC high-rising and low-rising tones in interlingual speech perception by naïve SC listeners. It is also found that the interlingual tonal identification is biased by the pitch height in the specific rendition of the word: when JM final rising words carry relatively lower ending pitch contours, they were more likely to be identified as the SC final low-rising pseudo-words. The pitch height in the first syllable also slightly affects the interlingual tonal perception, in that the higher the previous pitch is, the more likely the word is to be identified as the SC final low-rising pseudo-words.

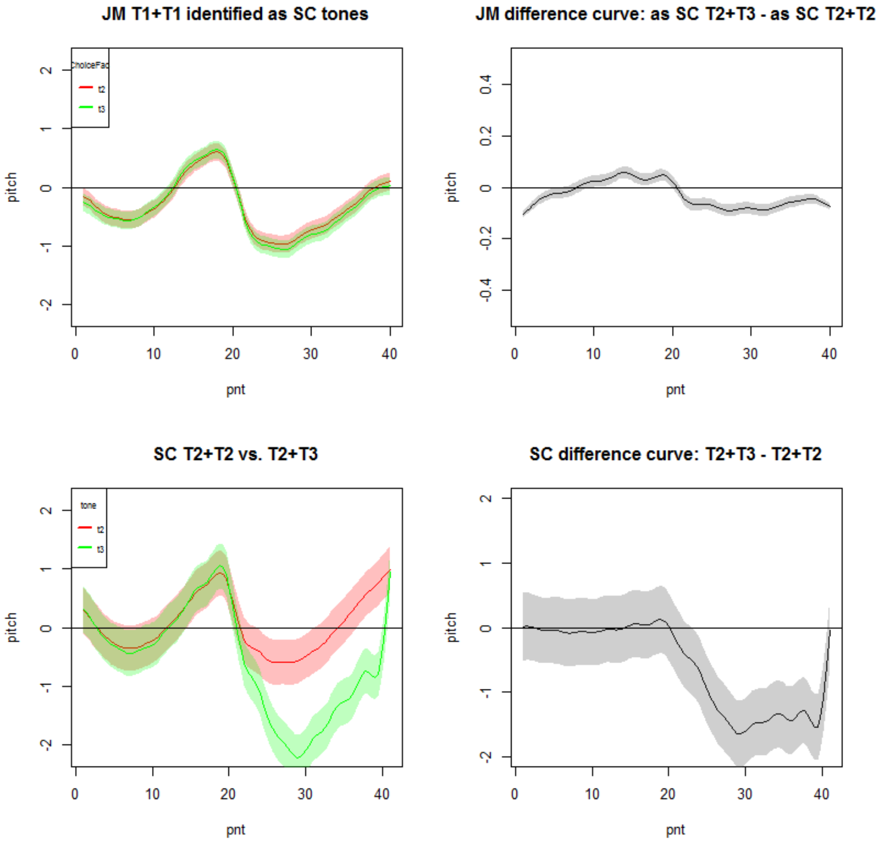

Figure 2. *Upper panel: estimated smoothes (left, random effects removed) and the corresponding difference curve (right) for JM final-rising words (JM T1+T1) identified as SC final-high-rising (SC T2+T2) and final-low-rising (SC T2+T3) pseudo-words. Lower panel: estimated smoothes (left) and the corresponding difference curve (right) for SC final-high-rising (SC T2+T2) and final-low-rising (SC T2+T3) pseudo-words.*

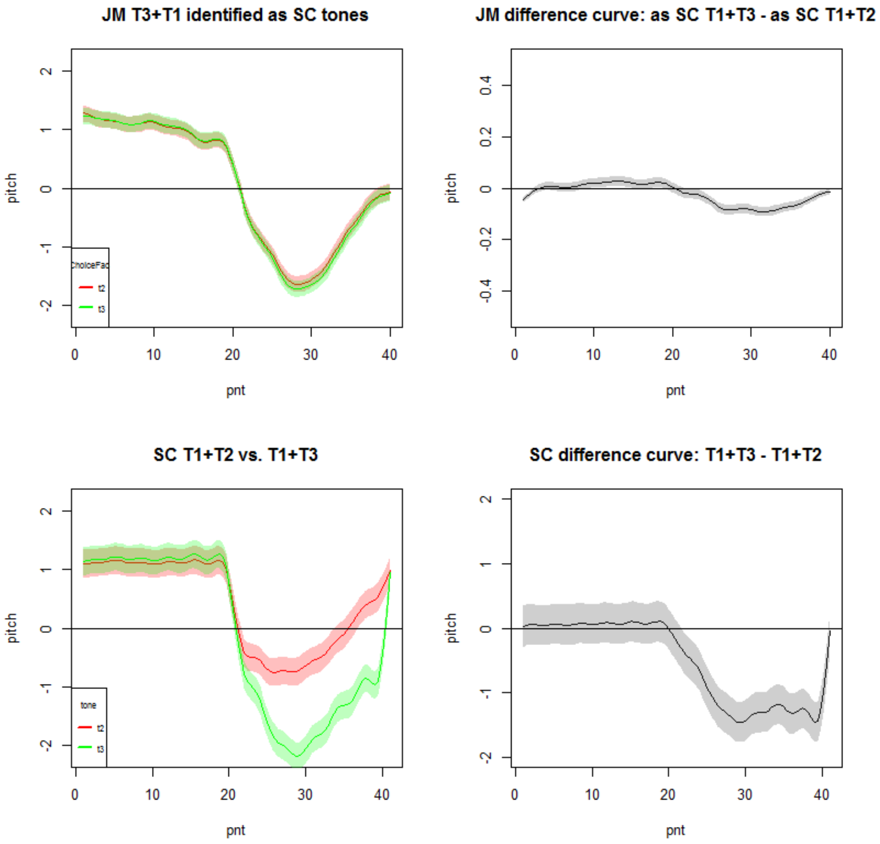

Figure 3. *Upper panel: estimated smoothes (left, random effects removed) and the corresponding difference curve (right) for JM final-rising words (JM T3+T1) identified as SC final-high-rising (SC T1+T2) and final-low-rising (SC T1+T3) pseudo-words. Lower panel: estimated smoothes (left) and the corresponding difference curve (right) for SC final-high-rising (SC T1+T2) and final-low-rising (SC T1+T3) pseudo-words.*

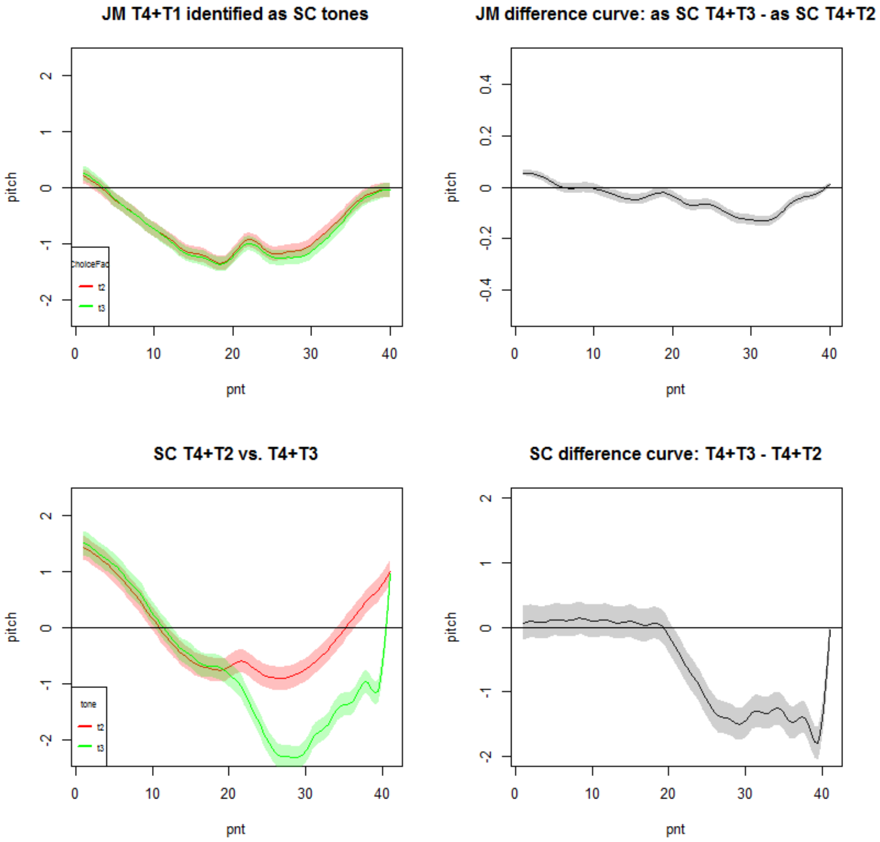

Figure 4. *Upper panel: estimated smoothes (left, random effects removed) and the corresponding difference curve (right) for JM final-rising words (JM T4+T1) identified as SC final-high-rising (SC T4+T2) and final-low-rising (SC T4+T3) pseudo-words. Lower panel: estimated smoothes (left) and the corresponding difference curve (right) for SC final-high-rising (SC T4+T2) and final-low-rising (SC T4+T3) pseudo-words.*

2.4 Experiment 3: bilingual semantic priming

Experiment 3 used the lexical decision task and semantic priming paradigm to investigate whether the two SC rising tones can both activate SC-JM tonal bilinguals' JM lexical representations, whether the mapping goodness of the two SC tones to the JM tonal categories affects the speed of JM lexical activation, and whether the mapping goodness also affects the semantic activation of the JM words.

2.4.1 Participants

Fifty-five native SC-JM tonal bilinguals from Jinan (15 male and 40 female, aged between 19 and 36 years, M = 23, SD = 3.85; 45 SC-dominant or balanced, 10 JM-dominant) participated in this experiment in exchange for payment. All participants were right-handed, received their literacy educations in SC, and learned some English at school. Four participants from the Jinan group and 15 participants from the Beijing group also had some knowledge of other non-tonal foreign languages, such as French and German.

2.4.2 Design and Stimuli

The present study adopted the semantic priming paradigm. Pairs of SC disyllabic pseudo-words were constructed with non-sense combinations of Chinese characters as primes, with one member ending with a high-rising tone and the other member ending with a low-rising tone. Each pair of pseudo-words was constructed so that their SC pronunciations may be mapped onto the same JM real words, which carry final rising tones. The matching JM words were selected from a 400-word corpus produced by 42 JM speakers (as used in Experiments 1 and 2). The two-to-one interlingual tonal mapping seems to cause no problem for the younger generations of SC-JM bilinguals.

To select semantically related targets, we combined the procedures used in earlier studies (Sumner & Samuel, 2005; Thierry & Wu, 2007). First a printed list of the primes was presented to a different group of 16 native SC speakers who were instructed to write down a related word for each item. One or two related targets were chosen for further selection based on the number of participants who wrote the given target as response. Then we formed word-pairs by crossing these potential targets with the primes and the semantic relatedness of these word pairs were rated on a scale from 1 to 5 by a group of 20 native SC speakers, including some of the above-mentioned native SC speakers and some new raters. One semantically related target and one semantically unrelated target were accordingly selected for each prime. For each JM real word, we used its related SC high-rising pseudo-word, its related SC low-rising pseudo-word, its semantically related target, and its semantically unrelated target as stimuli, as shown in the appendix together with the semantic relatedness data. The stimuli also included 32 JM non-words. A male native bilingual who is highly proficient in both dialects then produced these words and non-words accordingly in JM and SC (also a trained phonetician with Putonghua Proficiency Test Certificates- Level1B).

The design was a four-by-four Latin square design. Test sets were split into four lists, participants were also split into four groups, and the combination of prime conditions (final high-rising or low-rising) and target conditions (semantically related or semantically unrelated) were counterbalanced across the participants and test lists, so that each participant experienced every condition in the same number of trials and heard one prime and one target from each set.

2.4.3 Procedure

Participants were tested individually in a quiet room using the E-Prime software (Schneider et al., 2002). They were told that they would hear a series of sound sequences and that they were required to decide whether or not each of these sound sequences was a real word. Each item was played binaurally through headphones, with instructions on the screen. A new trial started 1,000 ms after the participant responded to the previous trial, or 1,000 ms after the response time exceeded 5 s. Each target was played directly after its corresponding prime with 1,000-ms ISI. The prime-target pairs were separated with non-word fillers and no prime-target pair was directly followed by another prime-target pair. In this way, we tried to control the potential phonological priming between the primes. The crucial test was preceded by a practice block including 5 words and 5 non-words.

2.4.4 Analysis and results

Analysis 1: Word-identification rates and reaction times to the primes. A *word-identification rate* was defined as the probability that a SC pseudo-word was identified by the SC-JM bilinguals as a real JM word. The average word-identification rates of both final high-rising and low-rising primes (78.18% and 67.95%, respectively) are above 50%. The identification data were collapsed into percentage scores indicating by-participant and by-item word-identification rate. Both by-participant and by-item t-test showed that the high-rising SC pseudo-words were more likely be identified as real JM words, by-participant t (54) = 3.38, $p < 0.01$, by-item t (15) = 2.98, $p < 0.01$.

Only the reaction times to the SC pseudo-words, which were accepted as real JM words, were taken into consideration. The reaction time data were also collapsed by participant and by item. Both by-participant and by-item t-test showed that the bilinguals responded faster to the high-rising SC pseudo-words than to the low-rising ones, by-participant t (54) = -5.49, $p < 0.001$, by-item t (15) = -5.20, $p < 0.001$.

These results indicate that both final high-rising and low-rising SC pseudo-words can be accepted as JM real words, although the final high-rising pseudo-words were more quickly and more likely to be accepted as real JM words.

Analysis 2: Accuracies and reaction times to the targets. The responses to the targets were influenced by more factors than the primes. We built Linear Mixed Effect (LME) models for the accuracy (ACC) and reaction time (RT) data (D. Bates, Maechler, Bolker, & Walker, 2013; R_Core_Team, 2013), including *Prime Condition* (final high-rising/final low-rising), *Target Condition* (related/unrelated), and their interactions as the fixed predictors. The candidates for the random terms included by-participant and by-target random intercepts and by-prime random intercepts nested under the related JM real word, as well as possible random slopes. The structure of the random terms in the models reported here was selected via model comparison based on likelihood ratio tests.

Logistic Linear Mixed Effect (LME) models were built for the binominal accuracy data. The selected random predictors were by-participant and by-target random intercepts, $X^2_{1|\text{participant}} = 7.44$, $p_{1|\text{participant}} < 0.001$, $X^2_{1|\text{target}} = 2.95$, $p_{1|\text{target}} < 0.1$. Parametric bootstraps (Singmann, 2014) showed that the main effect of *Target Condition* was significant, $F = 9.67$, $p < 0.01$. However, the main effect of *Prime Condition*, $F = 0.06$, n.s., and the interaction of *Prime Condition* and *Target Condition*, $F = 0.05$, n.s, were insignificant. Compared with the semantically unrelated targets, target $\text{ACC}_{\text{high-rising-prime}} = 96.4\%$, target $\text{ACC}_{\text{low-rising-prime}} = 96.4\%$, the semantically related targets showed significantly higher accuracy rates, target $\text{ACC}_{\text{high-rising prime}} = 99.5\%$, target $\text{ACC}_{\text{low-rising prime}} = 99.1\%$.

Linear Mixed Effect (LME) models were built for the reaction times to the targets. Only correct responses to the targets were considered and the reaction times were log-transformed to improve the distribution of the data. In the following analysis, a model was fit with all the data points, and then a model criticism removed the data points with standardized residuals exceeding 2.5 standard deviation units from the data set (less than 2.5% of the data) and refitted the model with the trimmed data set. We report the model statistics from the trimmed models, with Satterthwaite approximation for degrees of freedom (Kuznetsova, Brockhoff, & Christensen, 2013).

The first LME models were built for the responses collected after the corresponding primes were identified as real JM words. The selected random predictors were by-participant, by-related-JM real words, and by-target random intercepts, $X^2_{1|\text{participant}} = 124.64$, $p_{1|\text{participant}} < 0.001$, $X^2_{1|\text{related-JM real-word}} = 6.84$, $p_{1|\text{related-JM real-word}} < 0.01$, $X^2_{1|\text{target}} = 47.99$, $p_{1|\text{target}} < 0.001$. As shown in Figure 5, the main effect of *Target Condition* was significant, F (df = 543.73) = 59.04, $p < 0.001$. However, the main effect of *Prime Condition*, F (df = 537.36) = 0.61, n.s., and the interaction of *Prime Condition* and *Target Condition* was insignificant, F (df = 532.63) = 0.00, n.s. Compared to the semantically unrelated targets, the semantically related targets were processed faster.

Figure 5. *Interaction plot of Prime Condition (columns) and Target Condition (lines) for Experiment 3 – Analysis 2.*

The second LME models were built for the responses collected after the corresponding primes were rejected as non-words. The selected random predictors were by-participant and by-target random intercepts, $X^2_{1|\text{participant}} = 49.75$, $p_{1|\text{participant}} < 0.001$, $X^2_{1|\text{target}} = 61.10$, $p_{1|\text{target}} < 0.001$. However, none-of the fixed predictors was significant, $F_{\textit{Prime-Condition}}$ (df=174.13) = 0.29, n.s., $F_{\textit{Target-Condition}}$ (df= 190.10) = 0.98, n.s., $F_{\textit{Prime}\times\textit{Target}}$ (df= 170.37) = 1.54, n.s. The reaction time was slightly reduced for the semantically related targets primed by final high-rising pseudo-words, although the difference was insignificant.

In sum, both final high-rising and final low-rising SC pseudo-words improved the accuracy and reduced the reaction times of the targets which are semantically related to the corresponding JM real words of the primes. The semantic priming effect was only salient when the SC pseudo-word primes were accepted as real JM words by the bilinguals.

2.5 Discussion and conclusion

2.5.1 Main results and interpretation

Experiments 1 and 2 together clarified the acoustic distribution and perceptual status of the SC and JM rising tones. Experiment 1 verified the impression that the acoustic distribution of the JM rising tone largely overlaps with both the SC high-rising and low-rising tones. As shown in Figure 1, the overlap with the SC high-rising tone is larger than the overlap with the SC low-rising tone. The GAM modeling also showed that the distributional center of the JM rising tone lies between the SC high-rising and low-rising tones. Experiment 2 verified that SC native monolinguals can perceive the JM rising tone as either the SC high-rising or low-rising tone. The interlingual tonal identification is largely based on how high the pitch is in the specific rendition. JM final rising words were more likely to be identified as the SC final low-rising pseudo-words when the syllable carried a relatively lower pitch contour. We also found that the asymmetry is more pronounced in the interlingual perception of the JM low-falling + rising combination as SC falling + high-/low-rising combination. This is probably because the JM low-falling is lower than the SC falling tone, which not only caused the following JM rising to be perceived as relatively higher, but also caused the whole pitch contour to match the SC high-rising even better.

Experiment 3 first verified that both the SC high-rising and low-rising tones can be accepted as the JM rising tone in JM lexical access and activate JM lexical nodes. The final high-rising SC pseudo-words were more quickly and more likely to be accepted as real JM words. Thus, the asymmetry also persists in interlingual lexical access. However, the asymmetry is not kept in the corresponding semantic activation. As long as the prime was accepted as a JM real word, it primed the semantically related target but whether the prime carried high-rising or low-rising tone makes no difference.

2.5.2 Theoretical implication

The interlingual perception of the tonal two-to-one mapping is consistent with the previous findings based on vowels and consonants (Aoyama et al., 2004; Best & Strange, 1992; Bohn & Flege, 1990; Flege et al., 1997; Iverson et al., 2003). The asymmetry in two-to-one interlingual mapping not only exists in vowels (Bohn & Flege, 1990; Flege et al., 1997) and consonants (Aoyama et al., 2004; Iverson et al., 2003), but also exists in tones. The asymmetry in tonal mapping has acoustic basis and perceptual effects, just as in segmental mapping.

The pitch register of the previous syllable serves as a reference in interlingual tonal identification. This finding is consistent with the previous findings. The same physical stimuli can be perceived as higher when the previous stimuli are low in pitch. This applies to both acoustic pitch perception and monolingual tonal perception (Fox & Qi, 1990; Leather, 1983; Lin & Wang, 1984; Moore & Jongman, 1997; Wong & Diehl, 2003; Wu, 2011). Obviously this also applies to interlingual tonal perception.

Previous studies on vowels and consonants support that the L2 acoustic input is captured by the native phonemic categories in lexical access (A. Cutler et al., 2006; Weber & Cutler, 2004). Also, it has been shown that Greek–English early sequential bilinguals gave different consonantal category-goodness ratings for the same physical stimuli depending on the language mode (Antoniou et al., 2012). However, there is limited evidence in support of overlapping divisions of the tonal acoustic space according to the language mode in lexical access. The present study first verified that the acoustic distribution of the JM rising tone overlaps with that of both SC rising tones and then found that the canonical realizations of the SC high-rising and low-rising tones could both be accepted as the JM rising tone in lexical access. Thus, two physical stimuli belonging to different tonal categories in one dialect can be captured by the same tonal category in the other dialect. The SC-JM early bilinguals do not store the JM rising tone integrated with the more similar SC high-rising tone. Instead, our finding supports the claim that two tonal representations belonging to different tonal systems in the bilingual mind can be associated with highly overlapping acoustic distributions.

Effects of consonantal and vowel asymmetry have been shown in interlingual lexical access in previous studies (Escudero et al., 2008; Weber & Cutler, 2004, 2006). The current study showed that the tonal asymmetry also affects how the tonal representations activate lexical representations. Since the SC high-rising tone overlaps more with the JM rising tone in acoustic distribution and its distributional center is closer to that of the JM rising tone, it is not surprising to find that the final high-rising SC pseudo-words are more likely to activate the corresponding JM real words, and do so faster. The current study supported that the asymmetric interlingual lexical access also involves tonal categories.

Using a semantic priming paradigm, the current study further distinguished the effect of phonetic asymmetry in interlingual lexical activation and semantic activation. When priming SC pseudo-words with JM real words, the semantic priming effect is symmetric and not affected by the interlingual category-goodness. Although it is strongly supported in the previous studies that the phonological and semantic activations happen in parallel in speech comprehension (Grosjean, 1980; William Marslen-Wilson, 1984; W. D. Marslen-Wilson & Welsh, 1978; O'Rourke & Holcomb, 2002; Rodriguez-Fornells et al., 2002; Van Petten et al., 1999; Zwitserlood, 1989), the influence of interlingual category-goodness is kept in lexical access but lost in semantic activation. This finding suggests that there is some discreteness in the step between lexical activation and semantic activation.

To sum up, the current study investigated the two Standard Chinese rising tones and their corresponding Jinan Mandarin rising tones in speech production, perception, and lexical access. The three rising tones form the two-to-one interlingual mapping pattern. The JM rising tone overlaps with both SC rising tones in speech production and speech perception and the interlingual mapping is asymmetric. This asymmetry affects interlingual lexical access by SC-JM bilinguals and supports the theoretical claim that the tonal acoustic space can be divided in an overlapping way in bilingual lexical access. The acoustic input is captured by the native tonal categories depending on the language mode. However, the asymmetric mapping does not affect the semantic activation after lexical access, which suggests some discreteness in the step between lexical activation and semantic activation.

References

Antoniou, M., Tyler, M. D., & Best, C. T. (2012). Two ways to listen: Do L2-dominant bilinguals perceive stop voicing according to language mode? *Journal of Phonetics, 40*(4), 582-594. doi: 10.1016/j.wocn.2012.05.005

Aoyama, K., Flege, J. E., Guion, S. G., Akahane-Yamada, R., & Yamada, T. (2004). Perceived phonetic dissimilarity and L2 speech learning: The case of Japanese/r/and English/l/and/r. *Journal of Phonetics, 32*(2), 233-250.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2013). lme4: Linear mixed-effects models using Eigen and S4 (Version R package 1.0-4.). Retrieved from http://CRAN.Rproject.org/package=lme4.

Bates, D. M. (2010). lme4: Mixed-effects modeling with R. *URL* *http://lme4. r-forge. r-project. org/book.*

Best, C. T., & Strange, W. (1992). Effects of phonological and phonetic factors on cross-language perception of approximants. *Journal of Phonetics, 20*(3), 305-330.

Boersma, P., & Weenink, D. (2014). Praat: doing phonetics by computer [Computer program] Version 5.3.82 retrieved 26 July 2014 from http://www.praat.org/.

Bohn, O.-S., & Flege, J. E. (1990). Interlingual identification and the role of foreign language experience in L2 vowel perception. *Applied Psycholinguistics, 11*(03), 303-328. doi: 10.1017/S0142716400008912

Broersma, M., & Cutler, A. (2008). Phantom word activation in L2. *System, 36*(1), 22-34. doi: 10.1016/j.system.2007.11.003

Chandrasekaran, B., Krishnan, A., & Gandour, J. T. (2009). Relative influence of musical and linguistic experience on early cortical processing of pitch contours. *Brain and Language, 108*(1), 1-9. doi: 10.1016/j.bandl.2008.02.001

Chen, J. Y., Chen, T. M., & Dell, G. S. (2002). Word-form encoding in Mandarin Chinese as assessed by the implicit priming task. *Journal of Memory and Language, 46*(4), 751-781. doi: 10.1006/jmla.2001.2825

Cutler, A., & Otake, T. (2004). Pseudo-homophony in non-native listening. *The Journal of the Acoustical Society of America, 115*(5), 2392-2392. doi: 10.1121/1.4780547

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

Dufour, S., Nguyen, N., & Frauenfelder, U. H. (2007). The perception of phonemic contrasts in a non-native dialect. *The Journal of the Acoustical Society of America, 121*(4). doi: 10.1121/1.2710742

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

Flege, J. E., Bohn, O. S., & Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics, 25*, 437-470. doi: 10.1006/jpho.1997.0052

Fox, R., & Qi, Y. Y. (1990). Context effects in the perception of lexical tone. *Journal of Chinese Linguistics, 18*(2), 261-284.

Garding, E., Kratochvil, P., Svantesson, J. O., & Zhang, J. (1986). Tone 4 and Tone 3 discrimination in modern Standard Chinese. *Language and speech, 29*(3), 281. doi: 10.1177/002383098602900307

Grosjean, F. (1980). Spoken word recognition processes and the gating paradigm. *Perception & Psychophysics, 28*(4), 267-283. doi: 10.3758/BF03204386

Hallé, P. A., Chang, Y. C., & Best, C. T. (2004). Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners. *Journal of Phonetics, 32*(3), 395-421. doi: 10.1016/S0095-4470(03)00016-0

Holcomb, P. J., & Neville, H. J. (1991). Natural speech processing: An analysis using event-related brain potentials. *Psychobiology, 19*(4), 286-300. doi: 10.3758/BF03332082

Iverson, P., Kuhl, P. K., Akahane-Yamada, R., Diesch, E., Tohkura, Y., Kettermann, A. (2003). A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition, 87*(1), B47-B57. doi: 10.1016/S0010-0277(02)00198-1

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. (2013). lmerTest: tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package). Retrieved from http://cran.r-project.org/web/packages/lmerTest/lmerTest.pdf

Ladefoged, P. (2001). *A Course in Phonetics* (4th ed.): Heinle & Heinle Pub.

Leather, J. (1983). Speaker normalization in perception of lexical tone. *Journal of Phonetics, 11*(4), 373-382.

Liang, J., & van Heuven, V. J. (2004). Evidence for separate tonal and segmental tiers in the lexical specification of words: A case study of a brain-damaged Chinese speaker. *Brain and Language, 91*(3), 282-293. doi: 10.1016/j.bandl.2004.03.006

Lin, T., & Wang, W. (1984). Shengdiao ganzhi wenti [Perception of tones]. *Zhongguo Yuyan Xuebao [Bulletin of Chinese Linguistics], 2*, 59-69.

Malins, J. G., & Joanisse, M. F. (2010). The roles of tonal and segmental information in Mandarin spoken word recognition: An eyetracking study. *Journal of Memory and Language, 62*(4), 407-420. doi: 10.1016/j.jml.2010.02.004

Marslen-Wilson, W. (1973). Linguistic structure and speech shadowing at very short latencies. *Nature, 244*(5417), 522-523. doi: 10.1038/244522a0

Marslen-Wilson, W. (1984). Function and process in spoken word recognition-A tutorial review. *Attention and performance, 10*, 125-150.

Marslen-Wilson, W. D., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology, 10*(1), 29-63. doi: 10.1016/0010-0285(78)90018-X

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

Mitterer, H., Chen, Y., & Zhou, X. (2011). Phonological Abstraction in Processing Lexical - Tone Variation: Evidence From a Learning Paradigm. *Cognitive Science, 35*(1), 184-197. doi: 10.1111/j.1551-6709.2010.01140.x

Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A. M., Jenkins, J. J., & Fujimura, O. (1975). An effect of linguistic experience: The discrimination of [r] and [l] by native speakers of Japanese and English. *Perception and Psychophysics, 18*(5), 331-340. doi: 10.1111/j.1551-6709.2010.01140.x

Moore, C. B., & Jongman, A. (1997). Speaker normalization in the perception of Mandarin Chinese tones. *The Journal of the Acoustical Society of America, 102*, 1864. doi: 10.1121/1.420092

O'Rourke, T. B., & Holcomb, P. J. (2002). Electrophysiological evidence for the efficiency of spoken word processing. *Biological psychology, 60*(2), 121-150. doi: 10.1016/S0301-0511(02)00045-5

Pallier, C., Colomé, A., & Sebastián-Gallés, N. (2001). The influence of native-language phonology on lexical access: Exemplar-based versus abstract lexical entries. *Psychological Science, 12*(6), 445-449. doi: 10.1111/1467-9280.00383

Peng, S. H. (2000). Lexical versus ‘phonological’representations of Mandarin Sandhi tones *Papers in Laboratory Phonology V: Acquisition and the Lexicon (Vol 5)* (pp. 152-167).

Qian, Z.-Y. (Ed.). (1997). *‘Jinan Fangyan Cidian’ (Jinan Dialect Dictionary)*. Nanjing: Jiangsu Education Press.

R_Core_Team. (2013). R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, version 2.15.

Rodriguez-Fornells, A., Schmitt, B. M., Kutas, M., & Münte, T. F. (2002). Electrophysiological estimates of the time course of semantic and phonological encoding during listening and naming. *Neuropsychologia, 40*(7), 778-787. doi: 10.1016/S0028-3932(01)00188-9

Sakamoto, Y., & Ishiguro, M. (1986). *Akaike Information Criterion Statistics.* doi:10.1080/00401706.1989.10488538

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

Shen, X. S., & Lin, M. (1991). A perceptual study of Mandarin tones 2 and 3. *Language and speech, 34*(2), 145. doi: 10.1177/002383099103400202

Singmann, H. (2014). afex: Analysis of Factorial Experiments. R package version 0.10-113. http://CRAN.R-project.org/package=afex.

Strange, W., Bohn, O. S., Trent, S. A., & Nishi, K. (2004). Acoustic and perceptual similarity of North German and American English vowels. *The Journal of the Acoustical Society of America, 115*, 1791. doi: 10.1121/1.1687832

Sumner, M., & Samuel, A. G. (2005). Perception and representation of regular variation: The case of final/t. *Journal of Memory and Language, 52*(3), 322-338. doi: 10.1016/j.jml.2004.11.004

Thierry, G., & Wu, Y. J. (2007). Brain potentials reveal unconscious translation during foreign-language comprehension. *Proceedings of the National Academy of Sciences, 104*(30), 12530. doi: 10.1073/pnas.0609927104

Van Petten, C., Coulson, S., Rubin, S., Plante, E., & Parks, M. (1999). Time course of word identification and semantic integration in spoken language. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*(2), 394. doi: 10.1037/0278-7393.25.2.394

Vance, T. J. (1987). *An introduction to Japanese phonology*: State University of New York Press Albany.

Warner, N., Jongman, A., Sereno, J., & Kemps, R. (2004). Incomplete neutralization and other sub-phonemic durational differences in production and perception: Evidence from Dutch. *Journal of Phonetics, 32*(2), 251-276.

Weber, A., & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language, 50*(1), 1-25. doi: 10.1016/S0749-596X(03)00105-0

Weber, A., & Cutler, A. (2006). First-language phonotactics in second-language listening. *The Journal of the Acoustical Society of America, 119*, 597. doi: 10.1121/1.2141003

Wong, P., & Diehl, R. L. (2003). Perceptual normalization for inter-and intratalker variation in Cantonese level tones. *Journal of Speech, Language, and Hearing Research, 46*(2), 413. doi: 10.1044/1092-4388(2003/034)

Wood, S. (2006). *Generalized additive models: an introduction with R*: CRC Press.

Wood, S. N. (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology), 73*(1), 3-36. doi: 10.1111/j.1467-9868.2010.00749.x

Wu, J. (2011). *Formulating the identification of mandarin Tone2 and Tone3 in multi-dimensional spaces.* Paper presented at the International Congress of Phonetic Sciences (ICPhS 2011).

Xi, J., Zhang, L., Shu, H., Zhang, Y., & Li, P. (2010). Categorical perception of lexical tones in Chinese revealed by mismatch negativity. *Neuroscience, 170*(1), 223-231. doi: 10.1016/j.neuroscience.2010.06.077

Ye, Y., & Connine, C. M. (1999). Processing spoken Chinese: The role of tone information. *Language and Cognitive Processes, 14*(5-6), 609-630. doi: 10.1080/016909699386202

Yuan, J. H., & Chen, Y. Y. (2014). 3rd tone sandhi in standard Chinese: A corpus approach. *Journal of Chinese Linguistics, 42*(1), 218-237.

Zhang, Q., & Damian, M. (2009). The time course of segment and tone encoding in Chinese spoken production: an event-related potential study. *Neuroscience, 163*(1), 252-265. doi: 10.1016/j.neuroscience.2009.06.015

Zhang, Q., & Zhu, X. (2011). The temporal and spatial features of segmental and suprasegmental encoding during implicit picture naming: An event-related potential study. *Neuropsychologia*. doi: 10.1016/j.neuropsychologia.2011.09.040

Zwitserlood, P. (1989). The locus of the effects of sentential-semantic context in spoken-word processing. *Cognition, 32*(1), 25-64. doi: 10.1016/0010-0277(89)90013-9

Appendix. Crucial stimuli

Prime			related target		unrelated target	
JM real word	SC Hr pseudo-word	SC Lr pseudo-word	JM word	relatedness	JM word	relatedness
声音 sound /ʂəŋ in (R-R)/[a]	绳银 /ʂəŋ in (Hr-Hr)/	绳引 /ʂəŋ in (Hr-Lr)/	图像 video /tʰu ɕiɑŋ (Hl-Lf)/	3.95	结果 result /tɕiɛ kuo (L-Hl)[b] /	1.8
飞机 airplane /fei tɕi (R-R)/	肥集 /fei tɕi (Hr-Hr)/	肥挤 /fei tɕi (Hr-Lr)/	大炮 canon /ta pʰau (R-Lf)/	3.8	各位 each person /kɤ uei (R-Lf)/	1.65
签收 sign for /tɕʰiɛn ʂou (R-R)/	前熟 /tɕʰiɛn ʂou (Hr-Hr)/	前手 /tɕʰiɛn ʂou (Hr-Lr)/	快递 express delivery /kʰuai tʰi (R-Lf)/	4.45	馒头 steamed bun /man tʰou (R-H)[b]/	1.3
先驱 pioneer /ɕiɛn tɕʰy (R-R)/	咸渠 /ɕiɛn tɕʰy (Hr-Hr)/	咸取 /ɕiɛn tɕʰy (Hr-Lr)/	革命 revolution /kɤ miŋ (Hl-Lf)/	4.4	帮忙 help /paŋ maŋ (L-Hf)[b]/	1.15
回家 return home /xuəi tɕia (Hl-R)/	灰荚 /xuəi tɕia (Hl-Hr)/	灰甲 /xuəi tɕia (Hl-Lr)/	过年 Chinese New Year /kuo niɛn (Lf-Hf)/	4.3	领先 leading /liŋ ɕiɛn (Hl-R)/	1.2
流失 outflow /liou ʂʅ (Hl-R)/	溜十 /liou ʂʅ (Hl-Hr)/	溜史 /liou ʂʅ (Hl-Lr)/	水土 land and water /ʂui tʰu (Hl-Hl)/	4.45	快递 express delivery /kʰuai tʰi (R-Lf)/	1.5
无非 nothing more than /u fei (Hl-R)/	乌肥 /u fei (Hl-Hr)/	乌匪 /u fei (Hl-Lr)/	就是 just like /tɕiu ʂʅ (R-Lf)/	3.75	休息 rest /ɕiu ɕi (Lf-L)[b]/	1.1
有些 some /iou ɕiɛ (Hl-R)/	优鞋 /iou ɕiɛ (Hl-Hr)/	优写 /iou ɕiɛ (Hl-Lr)/	许多 many /ɕy tuo (Hl-R)/	3.45	技术 technique /tɕi ʂu (Lf-L)[b] /	1.95
大家 everyone /ta tɕia (Lf-R)/	大荚 /ta tɕia (F-Hr)/	大甲 /ta tɕia (F-Lr)/	各位 each person /kɤ uei (R-Lf)/	4.45	图像 image /tʰu ɕiɑŋ (Hl-Lf)/	1.25
拜托 please /pai tʰuo (Lf-R)/	拜驼 /pai tʰuo (F-Hr)/	拜妥 /pai tʰuo (F-Hr)	帮忙 help /paŋ maŋ (L-Hf)[b]/	4.65	革命 revolution /kɤ miŋ (Hl-Lf)/	1.05

Prime			related target		unrelated target	
JM real word	SC Hr pseudo-word	SC Lr pseudo-word	JM word	relatedness	JM word	relatedness
信息 information /ɕin ɕi (Lf-R)/	信席 /ɕin ɕi (F-Hr)/	信喜 /ɕin ɕi (F-Lr)/	技术 technique /tɕi ʂu (Lf-L)[b] /	4.45	就是 just like /tɕiu ʂʅ (R-Lf)/	1.5
肉汁（儿）gravy /ʐou tʂər (Lf-R)/	肉直（儿） /ʐou tʂər (F-Hr)/	肉纸（儿） /ʐou tʂər (F-Lr)/	馒头 steamed bun /man tʰou (R-H)[b]/	3.8	首先 first /ʂou ɕiɛn (Hl-R)/	1.05
到家 arrive home /tau tɕia (Lf-R)/	到荚 /tau tɕia (F-Hr)/	到甲 /tau tɕia (F-Lr)/	休息 rest /ɕiu ɕi (Lf-L)[b]/	4.35	许多 many /ɕy tuo (Hl-R)/	1.15
原因 reason /yæn in (Hl-R)/	冤银 /yæn in (Hl-Hr)/	冤引 /yæn in (Hl-Lr)/	结果 result /tɕiɛ kuo (L-Hl)[b] /	4.7	大炮 canon /ta pʰau (R-Lf)/	1.2
第一 first /ti i (Lf-R)/	第移 /ti i (F-Hr)/	第以 /ti i (F-Hr)/	首先 first /ʂou ɕiɛn (Hl-R)/	4.6	水土 land and water /ʂui tʰu (Hl-Hl)/	1.05
上风 windward /ʂaŋ fəŋ (Lf-R)/	尚缝 /ʂaŋ fəŋ (F-Hr)/	尚讽 /ʂaŋ fəŋ (F-Hr)/	领先 leading /liŋ ɕiɛn (Hl-R)/	3.7	过年 Chinese New Year /kuo niɛn (Lf-Hf)/	1

a. Abbreviation for tones. SC: Hl = High-level (1), Hr = High-rising (2), Lr = Low-rising (dip tone) (3), F = Falling (4). JM: R = Rising (1), Hf = High-falling (2), Hl = High-level (3), Lf = Low-falling (4), H = High (only in sandhi), L = Low (only in sandhi)

b. Tone Sandhi in the table: R+Hf -> L-Hf, R+Hl -> L-Hl, Lf+LF -> Lf-L/R-Lf, Hf+neutral -> R-H, R+neutral -> Lf-L