Tone and intonation processing: from ambiguous acoustic signal to linguistic representation

Liu, M.

Author: Liu, M.
Title: Tone and intonation processing: from ambiguous acoustic signal to linguistic representation
Issue Date: 2018-11-01

Chapter 5

Effects of cross-dialect phonological similarity in segment and tone on bi-dialectal auditory word recognition: Evidence from Xi'an Mandarin and Standard Chinese

Abstract

The present study examined if and how cross-dialect phonological similarity in segment and tone affects bi-dialectal listeners' lexical access during spoken word recognition. Balanced bi-dialectal speakers of Xi'an Mandarin and Standard Chinese took part in an auditory-auditory priming experiment with a generalized lexical decision task. The primes were monosyllabic homophones from either Xi'an Mandarin or Standard Chinese while the targets were disyllabic Xi'an Mandarin or Standard Chinese words. Primes and the first syllable of the target words overlapped in both segment and tone within a dialect (identical) or across two dialects (interdialectal homophones), or they overlapped in segment only within a dialect or across two dialects. In addition, a control condition was included where primes and targets shared neither tone nor segment. Results showed that Standard Chinese primes did not yield significant priming effects for within- or cross-dialect segment-only overlap targets. Standard Chinese primes did not produce significant priming effects for within-dialect identical targets either. However, they did yield significant inhibitory priming effects for cross-dialect homophone targets. This overall pattern was reversed for Xi'an Mandarin primes because these primes were not treated differently from their interdialectal homophonous primes in the current mixed dialect setting. These results suggest that cross-dialect phonological similarity in segment alone does not affect lexical access in bi-dialectal auditory word recognition while cross-dialect phonological similarity in both segment and tone poses a threat to the recognition system of bi-dialectal listeners. We conclude that tonal information plays a significant role in constraining word activation in bi-dialectal auditory word recognition.

Keywords: bi-dialect, Standard Chinese, Xi'an Mandarin, segment, tone, cross-dialect homophones, auditory word recognition

5.1 Introduction

Spoken word recognition is a key aspect of language comprehension. To arrive at the correct recognition of a spoken word, listeners constantly map the incoming speech signal onto possible lexical representations in their mental lexicon until the best-matched candidate is found (Luce & Pisoni, 1998; Marslen-Wilson, 1987; McClelland & Elman, 1986; Norris, 1994). This process usually takes place in just a few hundred milliseconds and seems effortless. However, when the incoming speech signal becomes ambiguous to the listener, for example when there are phonologically similar words to the intended target within one language system or across two language systems mastered by the listener, the recognition system can be challenged (Dijkstra, Grainger, & Van Heuven, 1999) and additional processing costs might be incurred. How an ambiguous speech input due to phonological similarities across two closely related linguistic systems affects spoken word recognition is the focus of interest in this study.

An extreme case of phonological similarity within a language system is homophony. Intralingual homophones have been found to be more difficult to process than non-homophone controls in tasks such as sentence verification (Coltheart, Avons, Masterson, & Laxon, 1991), semantic categorization (Van Orden, 1987), proofreading (Daneman & Stainton, 1991), eye movements (Daneman, Reingold, & Davidson, 1995; Jared, Levy, & Rayner, 1999) and lexical decision (Ferrand & Grainger, 2003; Newman, 2012; Pexman, Lupker, & Jared, 2001; Unsworth & Pexman, 2003). These results have been taken as evidence for parallel activation of multiple lexical candidates within a language.

Studies on homophones across two language systems are mainly concerned with bilingual word recognition. Interlingual homophones have been found to be processed much more slowly and less accurately compared to monolingual controls by bilingual speakers in tasks such as language-specific/generalized lexical decision (Dijkstra et al., 1999; Doctor & Klein, 1992; Lagrou, Hartsuiker, & Duyck, 2011; Nas, 1983), gating (Grosjean, 1988; Schulpen, Dijkstra, Schriefers, & Hasper, 2003) and semantic-relatedness decision (Luo, Johnson, & Gallo, 1998; Ota, Hartsuiker, & Haywood, 2009). This holds across both the visual (Dijkstra et al., 1999) and the auditory modalities (Lagrou et al., 2011). In cross-modal form priming tasks, bilinguals also showed longer visual lexical decision times for targets preceded by the auditory interlingual homophone primes than those by monolingual control primes (Schulpen et al., 2003). Overall, these studies suggest parallel activation of homophone candidates from both languages and an interference effect of cross-language phonological similarity on word recognition.

Most studies on phonological similarity across two language systems have been conducted in Indo-European languages with an alphabetic writing system, such as Dutch-English. While there are interlingual homophones across these languages in the visual domain, within the auditory domain many of them differ in subphonemic features in their actual pronunciation. This difference might partly account for the neglected role of phonology in bilingual studies (Dijkstra et al., 1999).

About 70% of the word languages are tonal languages (Yip, 2002). Some of these languages differentiate lexical meanings via pitch changes. For example, in Standard Chinese, the same segment *ma* can mean *mother*, *hemp*, *horse* and *scold* with different pitch contours, known as lexical tones. Little, however, do we know about how lexical tones constrain auditory word recognition in bilingual tonal language speakers, when interlingual phonological similarities can be due to overlap in segment and/or tone.

It is also worth noting that under the cover term Chinese, there are many language varieties spoken in China, which share an abundant number of homophones. The majority of speakers are proficient in at least two varieties: the national language Standard Chinese and their regional native dialect. Little attention has been paid to the possible co-activation of homophones in bi-dialectal word recognition. The question that arises is whether in bi-dialectal lexical processing, homophones co-activate and interfere, as in the bilingual situation. Furthermore, for tonal language speakers, what role does tone play in the activation and processing of bi-dialectal lexical representations during spoken word recognition?

Studies on the role of tone in lexical processing have mostly been conducted on Chinese varieties (mostly Standard Chinese and Cantonese) in a monolingual context. The general consensus is that tonal information is used in recognition

(Ching, 1985; Fox & Unkefer, 1985). However, in sublexical tasks such as homophone decision (Taft & Chen, 1992), vowel-tone monitoring (Ye & Connine, 1999), same-different judgment and non-word identification (Cutler & Chen, 1997), the processing of tonal information has been shown to be more error-prone than segmental information. It was therefore concluded that tone plays a weaker role in word recognition than segment at the pre-lexical processing stage. In contrast, an argument for the relatively strong role of tone at the lexical processing stage, comparable to that of the segment, has gained support from lexical tasks in an ERP study (Schirmer, Tang, Penney, Gunter, & Chen, 2005) and an eye-tracking study (Malins & Joanisse, 2010).

An auditory-auditory priming paradigm has also been adopted to investigate the role of tone and segment in word recognition at the lexical processing stage. Contradictory results were obtained as to whether tonal information constrains lexical activation. Lee (2007) investigated monosyllabic word recognition in Standard Chinese. Primes and targets overlapped only in segment (e.g., *lou3-lou2*), only in tone (e.g., *cang2-lou2*), both in segment and tone (e.g., *lou2-lou2*), or – in the baseline condition – neither in segment nor tone (e.g., *pan1-lou2*). Standard Chinese listeners were asked to make lexical decisions over the targets. A facilitatory priming effect was only found when primes and targets overlapped in both segment and tone. Segment-only overlap (minimal tone pair) or tone-only overlap did not produce any priming effect, comparable to the baseline condition. Lee (2007) interpreted the absence of priming in the minimal tone pair in Standard Chinese as the use of tonal information to constrain lexical activation. Sereno and Lee (2015), however, raised the concern that Lee (2007) did not control for the tonal similarity of the prime-target pairs. They conducted a follow-up study with balanced tonal distribution in the prime-target pairs and replicated the identity priming effect in Lee (2007) for the segment and tone overlap condition. In addition, they found a segment-only overlap facilitation effect, though smaller than the identity priming effect. Tone-only overlap, on the other hand, produced significant inhibition.

Given the conflicting results, more research is clearly needed to establish the role of lexical tone in auditory word recognition in Standard Chinese. It is also important to note that most speakers of Standard Chinese are bi-dialectal speakers. Existing studies have not controlled for participants' dialect background, which could be a potential cause of the different roles of tone and segment found in the literature. This study therefore set out to tap directly into their role(s) in bi-dialectal speakers' lexical processing. Specifically, we investigated the effect of cross-dialect phonological similarity in segment and tone on auditory word recognition in a bi-dialectal context using the auditory-auditory priming paradigm.

5.2 The present study

The two dialects we examined in this study are Standard Chinese (SC) and Xi'an Mandarin (XM). They both belong to the Mandarin family, which is the largest of the ten major Chinese dialect groups, following Chappell (2001)(but see Li & Thompson, 1981 which argues for a seven major dialect groups). SC is the most influential language within the Mandarin family. It is the official language of China and the medium of education. XM, on the other hand, is a local dialect spoken in the urban areas of Xi'an, the capital of Shaanxi Province. XM shares a common logographic writing system with SC and bears high resemblance with SC in terms of lexical items and syntactic forms. Moreover, XM exhibits large overlap of segmental features and tonal inventories with SC.

There are four lexical tones in SC, commonly known as the high-level (T1), mid-rising (T2), low-dipping (T3) and high-falling (T4) tones. XM also has four lexical tones, which are typically described as low-falling (T1), mid-rising (T2), high-falling (T3) and high-level (T4). Here, the terms T1-T4 are adopted to suggest that words which share the same tonal categories across the two dialects are etymologically-related translation equivalents in most cases. And this is consistently the case in our stimuli. Results of a tone production and a perception experiment (reported in Chapter 4) showed that tones with similar contours between the two dialects are basically perceived to be the same (see Table 1 for a summary). Specifically, tonal pairs of level contour (SC_T1 vs. XM_T4), rising contour (SC_T2 vs. XM_T2) and falling contour (SC_T4 vs. XM_T3) are perceived to be the same by the bi-dialectal speakers of XM and SC. Also, the tonal pair of low contour (SC_T3 vs. XM_T1) is generally perceived to be the same because XM_T1 sounds like an allotone of the citation form SC_T3. This systematic mapping of tones between SC and XM, together with the large overlap of segmental features between the two dialects, makes cross-dialect homophones prevalent in SC and XM. Such cross-dialect homophones exist in a consistent tonal mapping fashion. For example, with the segments being identical, a SC_T1 monosyllable (e.g., SC_T1, ma1/妈, "mother", high-level contour) is homophonous with a XM_T4 monosyllable (e.g., XM_T4, ma4/骂, "scold", high-level contour). Cross-dialect minimal tone pairs are also common in SC and XM. For example, a SC_T1 monosyllable (e.g., SC_T1, ma1/妈, "mother", high-level contour) shares the segmental structure but not the tonal contour with a XM_T1 monosyllable (e.g., XM_T1, ma1/妈, "mother", low-falling contour).

Table 1. *Paired tones with similar contours from Standard Chinese and Xi'an Mandarin.*

Tone pair	Standard Chinese (SC)			Xi'an Mandarin (XM)		
	Tonal category	Pitch value	Example	Tonal category	Pitch value	Example
Level contour	SC_T1	55	ma1/妈	XM_T4	55/44/45	ma4/骂
Rising contour	SC_T2	35	ma2/麻	XM_T2	24	ma2/麻
Low contour	SC_T3	214	ma3/马	XM_T1	21/31	ma1/妈
Falling contour	SC_T4	51	ma4/骂	XM_T3	52/53/42	ma3/马

Most monolingual priming studies on the role of segment and tone in auditory word recognition used monosyllabic primes and monosyllabic targets in their setup (Lee, 2007; Sereno & Lee, 2015; Yip, 2001). The present cross-dialect priming study instead used monosyllabic primes and disyllabic targets. This is because the dialect membership information is critical for lexical decision, and it is difficult for bi-dialectal listeners to determine whether a monosyllabic target belongs to SC or XM due to their great phonetic similarity. It was also possible to use disyllabic primes and disyllabic targets. However, Cutler and Chen (1995) have reported a positional effect of segmental and tonal similarity on disyllabic spoken word recognition in Cantonese. Segment and tone overlap in the first syllables produced inhibition (e.g., *ji6liu4-ji6liu5*/*to4fa1-to4foo1*), whereas segment and tone overlap in the second syllables generated facilitation (e.g., *to4va6-to2wa6*/*si6yip6-sue6yip6*). Given their findings, we chose monosyllabic primes and disyllabic targets to illuminate the pattern of bi-dialectal co-activation with the presentation of the prime and how the phonological similarities between the prime and target facilitate or inhibit target processing.

We constructed five types of relationships between the monosyllabic prime and the first syllable of the disyllabic target. They overlapped in both segment and tone within a dialect (identical) or across two dialects (interdialectal homophones). They overlapped in segment only within a dialect (within-dialect minimal tone pair) or across two dialects (cross-dialect minimal tone pair). The baseline condition was that they overlapped neither in tone nor segment within a dialect.

One factor that is known to influence bilingual lexical access is language proficiency (Jared & Kroll, 2001). Bilingual speakers often display high proficiency in one language over the other, with one language being more dominant. Previous cross-language phonological similarity studies have consistently reported an interference effect of phonology from the more dominant language to the less dominant language (e.g., Spivey & Marian, 1999, Marian & Spivey, 2003; Marian, Blumenfeld, & Boukrina, 2008), whereas there was mixed evidence of an interference effect of phonology (Lagrou et al., 2011) and a null effect (Weber & Cutler, 2004; Haigh & Jared, 2007) from the less dominant language to the more dominant language. We minimized language dominance influence by testing balanced, highly proficient bi-dialectal speakers of XM and SC. Furthermore, we adopted a generalized lexical decision task, similar to that in Doctor and Klein (1992), to ensure that the direction effect, if present, would not be biased by the target dialect of the task (Lemhöfer & Dijkstra, 2004). Since it can be difficult to control for the absolute proficiency level of bi-dialectals, we included primes and targets in both dialects by pairing all the target types with two prime types, one from SC (i.e., SC_T1), one from XM (i.e., XM_T4), which are cross-dialect homophones. In this way, if there is a language dominance effect, we are still able to detect it.

We expect that cross-dialect phonological similarity in both segment and tone would affect lexical access in bi-dialectal auditory word recognition. A facilitatory priming effect of the prime would be found not only for the within-dialect identical targets, but also for the cross-dialect homophone targets relative to the unrelated control targets. However, the prime might activate lexical candidates of homophones in both dialects, which could cause interference effects during the lexical decision of the within-dialect identical targets and cross-dialect homophone targets. The interference effect would in turn reduce the facilitatory priming effect of the primes for the within-dialect identical targets and for the cross-dialect homophone targets relative to the unrelated control targets. Consequently, the identity priming effect in the bi-dialect context might not be as strong as that in the monolingual context. Moreover, if tone plays a role in constraining lexical activation in word recognition, the segment-only overlap primes and targets (the minimal tone pair) would not prime each other. No priming effect would be found for either the within-dialect minimal tone pairs or cross-dialect minimal tone pairs. If tone does not play a role in constraining lexical activation in word recognition, the minimal tone pair would prime each other, but to a lesser extent than the priming effect between the identical prime and target. Since we tested balanced bi-dialectal listeners of XM and SC, we did not expect a direction difference in any priming effect. Any effect should be found in both directions, from SC to XM and from XM to SC.

5.3 Method

5.3.1 Participants

One-hundred balanced XM_SC bi-dialectal speakers (41 males, 59 females) were selected and paid to participate in the experiment. To assess their language proficiency in the two dialects, we asked participants to read the story “the North Wind and the Sun” in both SC and XM. In addition, an adapted version of the LEAP-Q questionnaire (Marian, Blumenfeld, & Kaushanskaya, 2007) was used to thoroughly check their language background and language proficiency. As can be seen in Table 2, all the selected participants were of high and comparable speaking proficiency (XM vs. SC: 7.8 vs. 8.1, $t(99) = -1.05$, $p = .30$) and spoken language comprehension skills (XM vs. SC: 8.3 vs. 8.6, $t(99) = -1.11$, $p = .27$) in the two dialects. They were born and raised in the urban areas of Xi'an and had no living experience outside of Xi'an. All were undergraduate or graduate students at local universities, with an age range from 19 to 28 ($M \pm SD$: 21.7 ± 3.2). None of them reported any speech or hearing disorders. Informed consent was obtained from all the participants before the experiment.

Table 2. *Self-ratings of language skills of Xi'an Mandarin and Standard Chinese by the participants.*

Measure	Xi'an Mandarin		Standard Chinese	
	M	*SD*	*M*	*SD*
Speaking proficiency	7.8	1.6	8.1	1.4
Level of spoken comprehension	8.3	1.5	8.6	1.0
Age onset speaking (in years)	4.0	4.5	5.6	9.4
Usage language interacting with family (in years)	16.5	8.1	8.3	10.2
Usage language in school/working settings (in years)	5.6	6.1	12.6	6.3
Percentage of current exposure time	29.2%	18.5%	58.9%	19.3%

Note. $N = 100$. Scale: 0 = None; 10 = Perfect.

Most of the selected participants learned to speak XM as the first dialect (D1) from their parents and then acquired SC from the age of 6 when they started to receive education at school. Some other participants learned to speak XM and SC almost simultaneously when they were young and had difficulty deciding whether SC or XM is their D1. Very few participants learned SC as D1 and started to speak XM afterwards. However, due to the fact that these participants have been exposed to XM-speaking settings by at least one of their parents from birth, they mastered XM as well as their SC despite starting to speak XM relatively later. All participants were thus selected as balanced XM_SC bi-dialectal speakers due to their language proficiency. For clarity purposes, we will not use the terminologies D1 and D2 but rather XM and SC in the present study. Overall, the average age at which participants began speaking XM and SC was 4.0 (*SD*: 4.5) and 5.6 (*SD*: 5.6), with the former being significantly younger than the latter ($t(99) = 5.09$, $p < .001$). It should also be noted that XM was mainly used for interacting with family ($M \pm SD$: 16.5 ± 8.1), whereas SC was mainly used in school/working settings ($M \pm SD$: 12.6 ±

6.3) across participants. At the time of testing, the percentage of time they were exposed to XM and SC in their daily life was 29.2% (*SD*: 18.5%) and 58.9% (*SD*: 19.3%), respectively. Apparently, despite their balanced linguistic competence in the two dialects and their relatively earlier exposure to XM, the participants had more access to SC compared to XM at the time of testing.

5.3.2 Stimuli

Forty SC_T1 monosyllables and their corresponding interdialectal homophonous XM_T4 monosyllables were selected as primes. Each monosyllabic prime was paired with five disyllabic targets, resulting in 400 prime-target trials (40 × 2 Prime types × 5 Target types = 400).

The first syllable of the disyllabic targets is our focus of interest. We therefore differentiated the five disyllabic target types according to property of their first syllable. The second syllable of the disyllabic targets always bears a T2, because T2 shows great resemblance in acoustic realization between the two dialects, i.e., SC_T2 maps onto XM_T2 both categorically and acoustically (see Table 1), and thus lends no ambiguity to the dialect membership of the disyllabic word by itself. T2 syllables, however, cue the dialect membership information of the disyllabic word together with the first syllable. For each prime type (e.g., SC_T1, "bang1/帮", "help"), the five disyllabic target types included a within-dialect segment and tone overlap target (Identical: e.g., SC_T1 target, "bang1mang2/帮忙", "help"), a within-dialect segment-only overlap target (D+Seg: e.g., SC_T4 target, "bang4qiu2/棒球", "baseball"), a cross-dialect segment-only overlap target (D−Seg: e.g., XM_T1 target, "bang1mang2/帮忙", "help"), a cross-dialect segment and tone overlap, i.e., an interlingual homophone, target (D−Homophone: e.g., XM_T4 target, "bang4qiu2/棒球", "baseball"), and a within-dialect control target which had neither segment nor tone overlap with the prime and served as baseline (Baseline: e.g., control target, "wan2cheng2/完成", "finish").

Note that the target words in the first four conditions were restricted to words that showed no pronunciation difference in the segment of the words between SC and XM. Also, there was no pronunciation difference between the segment of the prime and that of the first syllable of these targets. Regarding the control items, we ideally would have used words belonging exclusively to the dialect of the prime type. However, there are very few such words as the vocabularies largely overlap between the two dialects. What is possible is to narrow down the control words to those that show a pronunciation difference in terms of segment between the two dialects. An effort was made to obtain such control words for each prime type. Different control target words were selected for the SC primes and XM primes. Table 3 shows an example of the experimental design. The complete set of word stimuli is listed in Table D1 (see Appendix D).

Table 3. *An example of the experimental design for the prime-word target stimuli.*

	Prime type	Word target type				
		Identical	D+Seg	D−Seg	D−Homophone	Baseline
Tonal category	SC_T1	SC_T1	SC_T4	XM_T1	XM_T4	SC_T2
Tonal contour	level	level	falling	low	level	rising
Character	帮	帮忙	棒球	帮忙	棒球	完成
Pinyin	bang1	bang1mang2	bang4qiu2	bang1mang2	bang4qiu2	wan2cheng2
English	*help*	*help*	*baseball*	*help*	*baseball*	*finish*
Tonal category	XM_T4	XM_T4	XM_T1	SC_T4	SC_T1	XM_T2
Tonal contour	level	level	low	falling	level	rising
Character	棒	棒球	帮忙	棒球	帮忙	成熟
Pinyin	bang4	bang4qiu2	bang1mang2	bang4qiu2	bang1mang2	cheng2shu2
English	*stick*	*baseball*	*help*	*baseball*	*help*	*mature*

Note. We only listed the "Tonal category" and "Tonal contour" for the first syllable of the disyllables.

For all selected items, their listed frequency was controlled. Since there was no specific frequency reference available for XM, and XM shares most of its vocabulary with SC, we took the frequency of the translation equivalent in SC (e.g., SC_T4, "bang4qiu2/棒球", "baseball") as the frequency of the selected XM items (e.g., XM_T4, "bang4qiu2/棒球", "baseball"). In this study, the monosyllabic primes are frequent monosyllabic words with more than 4,500 occurrences in a corpus of 193 million words (Da, 2004). We ensured that each SC_T1 prime had a comparable word frequency to the matched XM_T4 prime

($t(39) = -1.85$, $p = .07$). The disyllabic targets were also of comparable frequency among conditions ($F(3, 156) = .52$, $p = .67$). According to the SUBTLEX-CH frequency list (Cai & Brysbaert, 2010), the average log10 word frequencies for the disyllabic targets were 2.45 (SC_T1/XM_T1), 2.34 (SC_T4/XM_T4), 2.46 (SC control) and 2.36 (XM control), respectively. None of the averaged word frequency was significantly different between target types.

In addition to the SC_T1 and XM_T4 monosyllabic primes and the disyllabic word targets, the same number of SC_T3 and XM_T1 monosyllabic primes and the disyllabic nonword targets were paired and added in the experiment (40 × 2 Prime types × 5 Target types = 400) based on similar logic. The SC_T3 and XM_T1 monosyllabic primes are legitimate monosyllables, and also interdialectal near-homophones, whereas the disyllabic nonword targets are not legitimate words in either SC or XM. The latter were made up of two permitted monosyllables, the second syllable of which bears a T2. For each prime type (e.g., SC_T3, "ma3/马"), the five disyllabic nonword target types included a within-dialect segment and tone overlap nonword target (Identical: e.g., SC_T3 target, "ma3duo2/马夺"), a within-dialect segment-only overlap nonword target (D+Seg: e.g., SC_T1 target, "ma1miao2/妈苗"), a cross-dialect segment-only overlap nonword target (D−Seg: e.g., XM_T3 target, "ma3duo2/马夺"), a cross-dialect segment and tone overlap, i.e., an interlingual homophone, nonword target (D−Homophone: e.g., XM_T1 target, "ma1miao2/妈苗"), and a within-dialect control nonword target which had neither segment nor tone overlap with the prime and served as baseline (Baseline: e.g., control target, "zhe4zuo2/这昨").

Again, we selected different control nonword targets for the SC_T3 primes and XM_T1 primes. The segment constraints between the primes and nonword targets were identical with those between the primes and word targets. Furthermore, none of the segments of the SC_T3 and XM_T1 primes and the nonword targets was ever used in the prime-word target conditions. Table 4 shows an example of the experimental design for the prime-nonword target stimuli. The complete set is listed in Table D2 (see Appendix D). The SC_T3 and XM_T1 monosyllabic primes are frequent monosyllabic words with more than 4,500 occurrences in a corpus of 193 million words (Da, 2004). Each

SC_T3 prime was ensured to have comparable word frequency to the matched XM_T1 prime ($t(39) = -0.52, p = .6$).

Table 4. *An example of the experimental design for the prime-nonword target stimuli.*

	Prime type	Nonword target type				
		Identical	D+Seg	D−Seg	D−Homophone	Baseline
Tonal category	SC_T3	SC_T3	SC_T1	XM_T3	XM_T1	SC_T4
Tonal contour	low-rising	low-rising	level	falling	low	falling
Character	马	马夺	妈苗	马夺	妈苗	这昨
Pinyin	ma3	ma3duo2	ma1miao2	ma3duo2	ma1miao2	zhe4zuo2
English	*help*	*N.A.*	*N.A.*	*N.A.*	*N.A.*	*N.A.*
Tonal category	XM_T1	XM_T1	XM_T3	SC_T1	SC_T3	XM_T3
Tonal contour	low	low	falling	level	low-rising	falling
Character	妈	妈苗	马夺	妈苗	马夺	厂闲
Pinyin	ma1	ma1miao2	ma3duo2	ma1miao2	ma3duo2	chang3xian2
English	*mother*	*N.A.*	*N.A.*	*N.A.*	*N.A.*	*N.A*

Note. We only listed the "Tonal category" and "Tonal contour" for the first syllable of the disyllables.

In total, there were 400 prime-word target trials and 400 prime-non-word target trials. Across all the items in the two dialects, the frequency of occurrence was comparable for the level-contour tones, low-contour tones and falling-contour tones.

5.3.3 Stimuli recording

A balanced XM_SC bi-dialectal male speaker was recruited to produce the stimuli in two separate blocks for the two dialects. This speaker was born and raised in the urban area of Xi'an and had no living experience out of Xi'an. He learned XM and SC simultaneously when he was young and was of high and comparable proficiency in the two dialects. He was an undergraduate student at a local university and used the two dialects equally frequently in his daily life. All the stimuli were recorded by him in a soundproof room at 16-bit resolution and a sampling rate of 44.1 kHz on a laptop via an external digitizer (UA-G1). The recorded stimuli were trimmed of silence and normalized amplitude for perception using Praat (Boersma & Weenink, 2015).

5.3.4 Procedure

We adopted an auditory-auditory priming paradigm in the perception experiment, with a monosyllabic prime preceding a disyllabic target in each trial. All the trials were distributed in a Latin Square design, so that participants only heard the same stimulus (prime or target) once during the experiment. Consequently, all the stimuli were divided into five lists. Each list contained both prime types and all the five word target types. For a given prime type, only one of the five target types occurred in every list. If a target (e.g., SC_T1 target, "bang1mang2/帮忙", "help") has already been assigned to the SC_T1 prime (e.g., SC_T1, "bang1/帮"), a different target (e.g., SC_T4 target, "bang4qiu2/棒球", "baseball") would be assigned to the corresponding XM_T4 prime (e.g., XM_T4, "bang4/棒") in the same list. The prime-nonword target trials were constructed in the same way in the list. In sum, each list included 80 prime-word target trials (40 × 2 Prime types × 1 Target type) and 80 prime-nonword target trials (40 × 2 Prime types × 1 Target type) with the five types of targets equally distributed.

Participants were tested individually on one list only in a soundproof booth of the behavioral lab at Shaanxi Normal University in Xi'an. Across all participants, the five lists were presented equally often (20 participants/list). All the trials in each list were presented to the participants using the E-Prime 2.0 software through headphones at a comfortable listening level. Trials were pseudo-randomized with the restriction that the shortest distance between the two interdialectal homophone primes was 9 trials and the shortest distance between two targets of the same type was 3 trials.

The experiment included a practice block and two experimental blocks. The practice block contained 10 trials to familiarize the participants with the task. These trials were not used in the experimental blocks. Each experimental block contained 80 trials. Between each block, there was a 3-minute break. Each trial started with a 100 ms warning beep, followed by a 300 ms pause. Participants then heard a pair of speech items separated by a 250 ms interval. The first item was a monosyllabic prime, and the second was a disyllabic target. Participants were asked to perform a generalized lexical decision task on the target as accurately and as quickly as possible, i.e., press the button labeled "yes" on the keyboard if the second item is either a SC word or a XM word, and press the button labeled "no" if the second item is neither a word in SC nor in XM. Button-press latencies were measured from the target offset. They were given up to 3 seconds after target offset to respond. Instructions were given both visually on screen in simplified Chinese characters and orally by the experimenter in mixed fashion of the two dialects (both SC and XM) before the experiment.

5.3.5 Data analysis

We restricted our analyses to the prime-word targets trials and discarded the prime-nonword target trials. The dependent variables included response accuracy and reaction time. Response accuracy was defined as the percentage of correct judgments of the word targets in the lexical decision task. Reaction time was defined as the response time relative to the offset of the word targets which were correctly responded to. To normalize the distribution, raw reaction times were transformed using the natural logarithm.

Statistical analyses were carried out with the package *lme4* (Bates, Mächler, Bolker, & Walker, 2015) in R version 3.1.2 (R Core Team, 2015). Analysis of response accuracy was performed using binomial logistic regression models, and analysis of reaction time was performed using linear mixed-effects regression models. The models included Prime type (SC_T1, XM_T4), Target type (Identical, D+Seg, D−Seg, D−Homophone, Baseline) and their interactions as fixed factors, and Subjects and Items as random factors. The fixed factors were added in a stepwise fashion and their effects on model fits were evaluated via model comparisons based on log-likelihood ratios. For Target type, all the conditions were first compared with the baseline condition. Post-hoc pairwise comparisons between different target conditions were conducted using *lsmeans* package (Lenth, 2016) with single-step *p*-value adjustment. For models of reaction time, trials with absolute standardized deviations exceeding 2.5 from the mean were considered as outliers and removed from further analysis.

5.4 Results

5.4.1 Response accuracy

Figure 1 presents the response accuracy for different target types preceded by SC_T1 primes and XM_T4 primes (see also Table D3 in Appendix D for details). Results showed a significant main effect of Target type ($\chi^2(4) = 18.73$, $p < .001$) and a significant two-way interaction of Prime type × Target type ($\chi^2(4) = 29.28, p < .001$). No main effect of Prime type was found ($\chi^2(1) = 0.40$, $p = .52$).

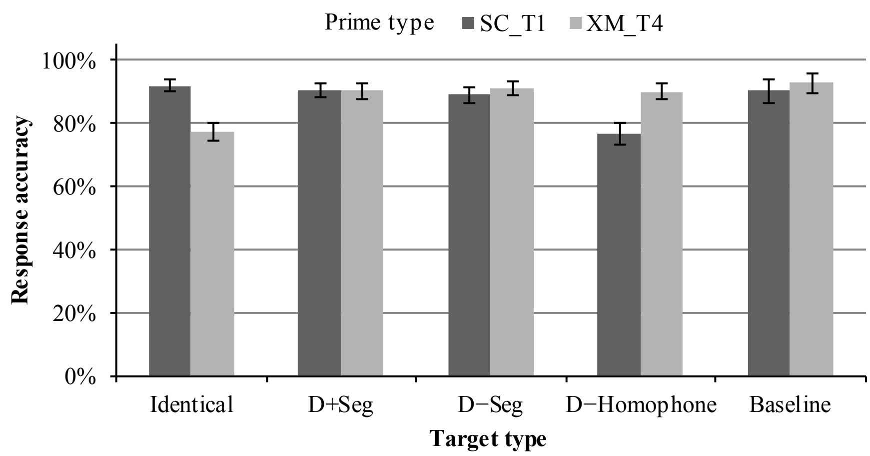

Figure 1. *Response accuracy for different target types preceded by SC_T1 primes (dark grey bars) and XM_T4 primes (light grey bars).*

Separate models were constructed for subset data of different prime types. When the prime was SC_T1, there was a significant main effect of Target type ($\chi^2(4) = 20.19$, $p < .001$), indicating that the response accuracy differed significantly among different target types. Further multiple pairwise comparisons showed that the response accuracy for the interdialectal homophone target was significantly lower than that for the other four target types (all *p*s $< .05$). No difference was found for any other pair of target types (all *p*s $> .05$). Overall, a SC_T1 prime made the recognition of its interdialectal homophone target XM_T4 words more erroneous for the balanced XM_SC bi-dialectal listeners.

When the prime was XM_T4, there was also a significant main effect of Target type ($\chi^2(4) = 27.88$, $p < .001$). Surprisingly, multiple pairwise comparisons showed that the response accuracy for the identical target, rather than the interdialectal homophone target from SC, was significantly lower than that for the other four conditions (all *p*s $< .05$). No difference was found for any other pair of conditions (all *p*s $> .05$). In other words, when a XM_T4 monosyllabic prime preceded a XM_T4 disyllabic target, the recognition of the latter became an error-prone process. Taken together, irrespective of whether the prime was the SC version or XM version of the homophone, the balanced XM_SC bi-dialectal listeners recognized the XM disyllabic target less accurately than the other target types.

5.4.2 Reaction time

2.3% of the data points were identified as outliers and removed from further analysis. Figure 2 presents the average reaction time for different target types preceded by SC_T1 primes and XM_T4 primes (see also Table D3 in Appendix D for details). The overall analyses showed a significant main effect of Target type ($\chi^2(4) = 15.22$, $p = .004$) and a significant two-way interaction of Prime type × Target type ($\chi^2(4) = 75.29$, $p < .001$). No main effect of Prime type was found ($\chi^2(1) = 0.09$, $p = .77$).

Separate models were constructed for subset data of different prime types. When the prime was SC_T1, there was a significant main effect of Target type ($\chi^2(4) = 46.05$, $p < .001$). Pairwise comparisons showed a non-significant facilitatory priming effect by the SC_T1 prime for the identical target compared to the baseline target ($\beta = -0.07$, $t = -1.31$, $p = .69$). Likewise, there was a non-significant facilitatory priming effect by the SC_T1 prime for the within-dialect segment-only overlap target (D+Seg) compared to the baseline target ($\beta = -0.07$, $t = -1.33$, $p = .67$). While it was quite unexpected that the identity priming and the within-dialect segment alone overlap priming did not reach significance, the priming trend for these two target types was indeed consistent with the priming effects reported in Sereno and Lee (2015). In addition, a comparison between the reaction time of these two target types showed almost no difference ($\beta = -0.001$, $t = -0.02$, $p = 1.00$), which is in contrast with the previous finding that the priming effect for the segmental and tonal overlap (identical) prime and target was stronger than that for the segment-only overlap prime and target (Sereno & Lee, 2015).

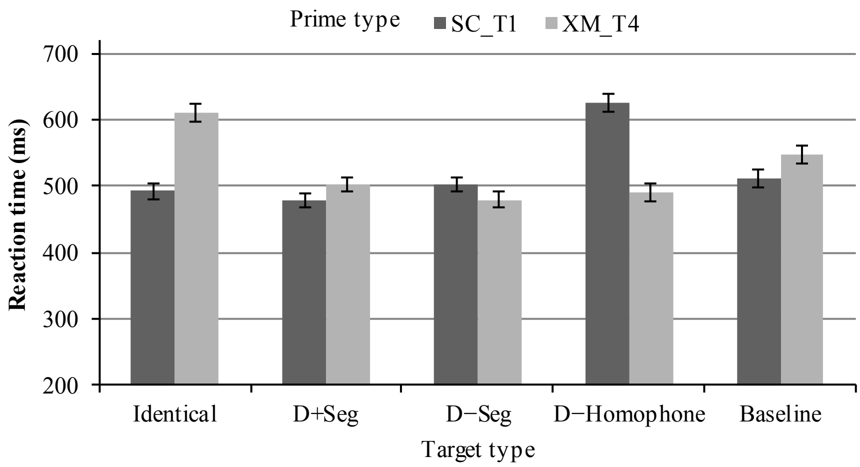

Figure 2. *Average reaction time for different target types preceded by SC_T1 primes (dark grey bars) and XM_T4 primes (light grey bars). The error bars represent the 95% confidence interval of the means across participants.*

As for the cross-dialect segment-only overlap target (D−Seg), its reaction time was comparable to that of the baseline target. No priming effect by the SC_T1 prime was therefore found ($\beta = 0.03$, $t = -0.61$, $p = .97$), seemingly indicating that cross-dialect segment-only overlap between primes and targets was not enough to yield priming between two tonal languages.

Regarding the interdialectal homophone target (D−Homophone), a significant inhibitory priming effect by the SC_T1 prime was found for this condition compared to the baseline target ($\beta = 0.25$, $t = 4.80$, $p < .001$). In fact, the reaction time for the interdialectal homophone target was not only longer than that for the baseline target, but also longer than that for all of the other target types (all *p*s < .001). The SC_T1 monosyllabic primes considerably slowed down the recognition of the interdialectal homophonous XM_T4 disyllabic targets.

Similar analyses were conducted for the data of XM_T4 primes, where a reversed pattern of reaction time was found compared to the SC_T1 prime. There was a significant main effect of Target type ($\chi^2(4) = 44.22$, $p < .001$). Contrary to the facilitatory priming trend of the SC_T1 prime for the identical target, the XM_T4 prime showed a significant inhibitory priming effect for the identical target ($\beta = 0.15$, $t = 2.85$, $p = .04$). The reaction time for the identical target was not only longer than that for the baseline target, it was also longer than that for all the other target types (all *p*s < .001). Overall, the XM_T4 monosyllabic primes considerably slowed down the recognition of the identical XM_T4 disyllabic targets. This result was rather surprising, considering that a complete overlap of both segment and tone between primes and targets (identical) has almost always shown a facilitatory priming effect (see Lee, 2007; Sereno & Lee, 2015).

The comparison between the reaction times of the within-dialect segment-only overlap targets and the baseline targets was carried out next. No priming effect by the XM_T4 prime was found for the within-dialect segment-only overlap target compared to the baseline target ($\beta = -0.07$, $t = -1.41$, $p = .62$).

Next was the comparison between the reaction times of the cross-dialect segment-only overlap targets and that of the baseline targets. There was a significant facilitatory priming effect of the XM_T4 prime for the cross-dialect segment-only overlap target relative to for the baseline target ($\beta = -0.16$, $t = -3.14$, $p = .02$).

Lastly, we compared the reaction time of the interdialectal homophone target and that of the baseline target following the XM_T4 primes. The former was significantly shorter than the latter ($\beta = -0.15$, $t = -2.98$, $p = .03$), suggesting an evident facilitatory priming effect of the XM_T4 prime for the interdialectal homophone target. Different from the results in the SC_T1 prime data showing that the facilitatory priming trend was found when the primes and targets belonged to the same dialect (i.e., in the SC_T1 prime-Identical target condition and the SC_T1 prime-D+Seg target condition), in the XM_T4 prime data we found a facilitatory priming effect when primes and targets belonged to different dialects (i.e., in the XM_T4 prime-D−Seg target condition and the XM_T4 prime-D−homophone target condition). A comparison between the reaction time of the interdialectal homophone targets and that of the cross-dialect segment-only overlap targets barely showed any difference ($\beta = 0.01$, $t = 0.16$, $p = 1.00$).

So far, we have found that when the prime was SC_T1, there was a similar facilitatory priming trend for the identical target and for the within-dialect segment-only overlap target. This facilitatory priming trend, however, did not hold for the cross-dialect segment-only overlap target. If the target was an interdialectal homophone of the prime, a significant inhibitory priming effect emerged from the prime to the target. The overall pattern was reversed for the XM_T4 prime data. When the prime was XM_T4, there was a significant inhibitory priming effect for the identical targets and a null effect for the within-dialect segment-only overlap targets. The cross-dialect segment-only overlap targets and the interdialectal homophone targets, on the other hand, showed similar facilitatory priming effects by the XM_T4 prime.

The pattern for the XM_T4 prime data was counterintuitive. If the SC_T1 prime and XM_T4 prime were represented equally well in the mental lexicon of the XM_SC bi-dialectal speakers, we would expect that the two prime types performed similarly on each target type. Yet the SC_T1 prime and XM_T4 prime behaved in a reversed fashion on different target types. What could be the possible reason for this?

One alternative way of viewing the reversed pattern of the XM_T4 prime data is that it could be a rearrangement of the pattern of the SC_T1 prime data, which led us to make the assumption that the XM_T4 prime might not be taken as XM_T4 itself, but as its interdialectal homophone counterpart, i.e. SC_T1. Since the two prime types were interdialectal homophones with almost no pronunciation difference, it is very likely that participants did not recognize the dialect membership of the XM_T4 prime and treated it as the SC_T1 prime. To test the validity of this assumption, we replotted Figure 2 according to the tonal category of the targets rather than the prime-target relationship. This is possible because the two prime types had targets of the same tonal categories which corresponded to different target types based on the prime-target relationship as stated earlier. The replotted results are presented in Figure 3. Just as what we have assumed, the pattern of the XM_T4 prime was very similar to that of the SC_T1 prime.

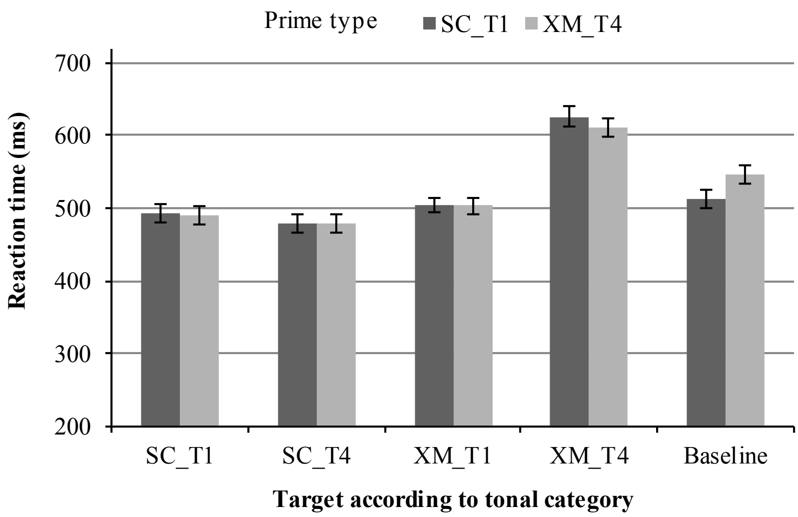

Figure 3. *Average reaction time for different targets arranged according to tonal category preceded by SC_T1 primes (dark grey bars) and XM_T4 primes (light grey bars). The error bars represent the 95% confidence interval of the means across participants.*

We also ran statistical analyses for the rearranged data. Linear mixed-effects regression models were built for reaction time with the fixed factors Prime type (SC_T1, XM_T4) and Target tonal category (SC_T1, SC_T4, XM_T1, XM_T4, Baseline). We only found a significant main effect of Target tonal category ($\chi^2(4) = 88.16, p < .001$). Neither a main effect of Prime type ($\chi^2(1) = 0.06, p = .81$) nor a significant two-way interaction of Prime type × Target tonal category ($\chi^2(4) = 2.85, p = .58$) was found. The null effect of the Prime type showed that the SC_T1 prime and XM_T4 prime were not treated differently. Since the pattern of the SC_T1 prime was more in line with previous results, we are tempted to claim that the XM_T4 monosyllabic primes were treated as their interdialectal homophonous SC_T1 equivalents in the current mixed dialect context.

5.5 General discussion

The present study investigated if and how cross-dialect phonological similarity in segment and tone affects lexical access in bi-dialectal auditory word recognition of balanced bi-dialectal tonal language listeners. In an auditory-auditory priming experiment with a generalized lexical decision task, we found that when the prime was in Standard Chinese, there was a non-significant within-dialect facilitatory priming trend for targets overlapping in both segment and tone with the prime, and also for targets overlapping only in segment with the prime. Both priming trends were of similar magnitude. The Standard Chinese prime did not produce any cross-dialect priming effect for the cross-dialect segment-only overlap target. It, however, produced a significant inhibitory priming effect, as evidenced by the lower response accuracy and longer reaction time, for the interdialectal homophone target relative to the unrelated Standard Chinese control target. The overall pattern was reversed when the prime was in Xi'an Mandarin, because the Xi'an Mandarin prime was treated as its interdialectal homophonous Standard Chinese prime in the current mixed dialect setting. It seems that cross-dialect phonological similarity in segment alone does not affect lexical access in bi-dialectal auditory word recognition while a cross-dialect phonological similarity in both segment and tone (cross-dialect homophones) does pose a threat to the recognition system of the bi-dialectal tonal language listeners. Tonal information plays a significant role in constraining word activation in bi-dialectal auditory word recognition.

The present study extended the investigation of the role of segment and tone in Standard Chinese auditory word recognition from a monolingual context to a bi-dialectal context. In the monolingual Standard Chinese context, a significant facilitatory priming effect has been consistently found for the identical primes and targets which overlap in both segment and tone. Also, a complete overlap in segment and tone between the prime and target has always shown more facilitation than segment-only overlap between primes and targets (Lee, 2007; Sereno & Lee, 2015). In the bi-dialectal context, we found that a complete overlap in segment and tone between the Standard Chinese primes and targets showed a non-significant facilitatory priming trend, and so did the Standard Chinese primes and targets overlapping in segment only. No magnitude difference was found between the priming trends of the two conditions. Our results contrast with the results in the monolingual context. The identity priming effect of Standard Chinese primes and targets in the current bi-dialectal context was not as strong as that in the monolingual context. It shrunk in size and did not reach significance, but the facilitatory priming trend was maintained. This result was in line with our hypothesis. As has been shown in previous bilingual studies, a homophone representation activated lexical candidates from both languages in a bilingual context (Schulpen et al., 2003; Lagrou et al., 2011). In the current bi-dialectal context, the presentation of the Standard Chinese version of the homophone prime seemed to have activated both the lexical candidates from Standard Chinese and its homophonous lexical candidates from Xi'an Mandarin, which caused an interference effect and reduced the identity facilitatory priming effect. The identical Standard Chinese primes and targets sharing phonological similarity in both segment and tone showed an overall facilitatory priming trend, despite the fact that the facilitation was reduced by the interference effect resulting from the coactivation of lexical candidates from both dialects by the Standard Chinese prime. A similar result was expected for primes and targets which share cross-dialect phonological similarity in both segment and tone (i.e., cross-dialect homophones). However, a significant inhibitory priming effect was found for the latter. The Standard Chinese prime considerably slowed down the recognition of interdialectal homophone targets with significantly more errors than the unrelated control target. As inhibitory priming has generally been taken as evidence of competition between lexical candidates activated by the prime and the target (Dufour & Peereman, 2003; Radeau, Morais, & Dewier, 1989; Slowiaczek & Hamburger, 1992), it seems that there was competition among the lexical candidates activated by the SC_T1 prime and the XM_T4 target. The inhibitory priming effects might have taken place under one of the following two scenarios. The first scenario is that the input signal of the auditory monosyllabic SC_T1 prime only activated the lexical representations of SC_T1, the presence of the XM_T4 target contrasts with the activated SC_T1 prime in dialect membership and the bi-dialectal listeners had to make a switch to activate the XM_T4 target to perform the task. Since dialect membership information can only be determined by the second syllable of the target together with the first syllable, the participants could not switch to Xi'an Mandarin until the presence of the second syllable, which caused longer reaction times and lower response accuracy in lexical decisions on XM_T4 target. The second scenario is that the input signal of the auditory monosyllabic SC_T1 prime activated the lexical representations of both SC_T1 and XM_T4, but with a stronger activation of SC_T1 and a weaker spreading activation of XM_T4. The coactivation of the SC_T1 and XM_T4 lexical representations caused an interference effect when the XM_SC bi-dialectal listeners made lexical decisions on the XM_T4 target, and the interference from SC_T1 would be much stronger than that from XM_T4, resulting in strong inhibition of the former to activate the latter. The lexical decision for the XM_T4 target was thus more time-consuming and error-prone compared to that for other target types. The first scenario advocates for a selective activation of Standard Chinese lexical representations by the SC_T1 prime, whereas the second scenario speaks for a non-selective parallel activation of both Standard Chinese and Xi'an Mandarin lexical representations with a stronger activation of the former than the latter by the SC_T1 prime. Both scenarios seem plausible to account for the inhibitory priming effect of the Standard Chinese prime for the Xi'an Mandarin homophone target. However, only the second scenario could account for the different priming results in the current bi-dialectal context with those in the monolingual context when the prime and target were both from Standard Chinese. Therefore, it is more likely that the SC_T1 prime activated the lexical representations of both Standard Chinese and Xi'an Mandarin with the former being more strongly activated than the latter.

A Standard Chinese prime produced an inhibitory priming effect for the cross-dialect segment and tone overlap target. It, however, did not produce any priming effect for the cross-dialect segment-only overlap target. Recall that there was also no significant priming effect of the Standard Chinese prime for the within-dialect segment-only overlap target. This lack of significant facilitatory priming in the minimal tone pair within and across dialects suggests that the members of the minimal tone pair were not treated as homophones. Tonal information was indeed used to constrain lexical activation in spoken word recognition, as claimed in Lee (2007). This finding echoes the previous findings that pitch accent in Japanese (Cutler & Otake, 1999) and stress in English (Cooper, Cutler, & Wales, 2002) and Dutch (Cutler & Van Donselaar, 2001) could be used to constrain lexical activations, together indicating that prosodic information might be universally adopted to constrain lexical activation in spoken word recognition. It should also be noted that the priming results for the within-dialect and cross-dialect minimal tone pair were not entirely the same. The Standard Chinese prime made it relatively easier (reflected in the shorter reaction time) for the XM_SC bi-dialectal listeners to recognize the Standard Chinese segment-only overlap targets than the Xi'an Mandarin segment-only overlap targets, though the reaction time of the two target types was not statistically different. The weaker effect of the prime for the cross-dialect segment-only overlap target than for the within-dialect segment-only overlap target again seems to be indicative of a weaker activation of the XM than SC representations by the Standard Chinese prime in the minds of the balanced XM_SC bi-dialectal listeners.

The above discussions were all for targets preceded by the SC_T1 primes. We also investigated the priming effects for all the target types preceded by the XM_T4 primes, which are homophonous with SC_T1. It was found that the overall pattern was reversed for all the target types with the XM_T4 primes compared to with the SC_T1 primes. Further analyses have shown that the Xi'an Mandarin primes were not treated differently from interdialectal homophonous SC_T1 primes in the current mixed dialect setting. The two primes were actually represented as the same in the minds of the balanced XM_SC bi-dialectal listeners. Like the SC_T1 prime, the XM_T4 prime activated the lexical representations of both SC_T1 and XM_T4, but with a stronger activation of SC_T1 and a weaker spreading activation of XM_T4 representations.

The fact that the SC_T1 prime and XM_T4 prime were not treated differently indicates that the balanced XM_SC bi-dialectal listeners did not perceive any subphonemic difference between the two primes. Both homophone primes were more strongly associated with the Standard Chinese representations than the Xi'an Mandarin representations under the current mixed dialect context. For balanced XM_SC bi-dialectal listeners, one might expect that both dialects were activated to similar degrees. The resulting stronger activation of the Standard Chinese representations relative to the Xi'an Mandarin representations by the homophone prime in the present study indicates that the two dialects were not represented to similar degrees in the minds of these XM_SC bi-dialectal listeners. In bilingual studies, Grosjean (1988; 1997) proposed that in the bilingual language mode, bilingual speakers choose a base language (the main language for communication) and call upon the other language (guest language) when necessary. The present study seems to show similar mechanisms for bi-dialectal word recognition in the bi-dialectal mode. Based on the stronger activation of Standard Chinese presentations than Xi'an Mandarin representations, the bi-dialectal listeners here appeared to choose Standard Chinese as their base language. This could possibly reflect that though the XM_SC bi-dialectal listeners we recruited were comparable in their language competence of the two dialects (see Table 2), perhaps they were not balanced XM_SC bi-dialectal listeners after all. They more likely had overall language dominance in Standard Chinese rather than in Xi'an Mandarin. One might argue that the balanced XM_SC bi-dialectal listeners we investigated were a mixture of bi-dialectal speakers with different orders of learning of the two dialects. It may be unfair to talk about language dominance for such a mixed group of participants without concerning their order of learning of the two dialects, as they might have performed differently when processing the two dialects. We admit that better control of the language background of the participants should have been made. However, the different orders of learning of the two dialects could hardly be a factor which has affected the priming pattern. A closer analysis showed that the bi-dialectal listeners with Standard Chinese as their D1 did not behave differently from those with Xi'an Mandarin as their D1. All the bi-dialectal listeners tended to be Standard Chinese dominant regardless of their D1. This is understandable given that Standard Chinese is the medium of education and it is more frequently used in campus life than Xi'an Mandarin by the bi-dialectal listeners.

In the field of spoken word recognition, previous cross-language phonological similarity studies are mostly concerned with bilingual word recognition in two Indo-European languages. The present study shifted the focus to cross-dialect phonological similarity effects in two tonal dialects. The results extend our understanding of the role of segment and tone in auditory word recognition in tonal languages from the monolingual context to the bi-dialectal context. The investigation of the cross-dialect homophone effect in bi-dialectal auditory word recognition also allows us to compare the effect of cross-dialect phonological similarity on bi-dialectal auditory word recognition with the effect of cross-language phonological similarity on bilingual auditory word recognition. In bilingual auditory word recognition, it has been found that the presentation of a homophone prime activated homophone representations from both languages; homophone primes of both languages facilitated the recognition of the L2 visual targets, yet there was competition between the two interlingual homophone representations compared to the monolingual control prime and target words (Schulpen et al., 2003). In the current bi-dialectal auditory word recognition, homophone primes of both dialects yielded a stronger activation of the lexical representations in Standard Chinese and a much weaker activation of the lexical representations in Xi'an Mandarin. There was intense competition of the activated representations of the homophone with the Xi'an Mandarin target, resulting in a significant inhibitory priming effect of the homophone prime for the Xi'an Mandarin target. Overall, the coactivation of the lexical representations in both dialects by the homophone prime in bi-dialectal auditory word recognition is in line with the coactivation of the lexical representations in both languages by the homophone prime in bilingual auditory word recognition, suggesting a non-selective processing mechanism in both bilingual lexical access and bi-dialectal lexical access during auditory word recognition.

The priming results in our bi-dialectal study also differed from that in the bilingual studies. For example, in the bilingual study, the homophone prime showed facilitatory priming for the L2 target (Schulpen et al., 2003), whereas in our bi-dialectal study, the homophone prime showed inhibitory priming for the Xi'an Mandarin target. This difference presumably results from the different strength of activation of the lexical representations in each language system due to difference in tasks, participants, and stimulus features between the two studies. More bi-dialectal research is clearly needed to understand the similarities and differences between bilingual and bi-dialectal word recognition.

5.6 Conclusion

To conclude, the present study showed that phonological similarity in segment alone did not affect lexical access in bi-dialectal auditory word recognition, whereas phonological similarity in both segment and tone (cross-dialect homophones) posed a threat to the recognition system of the bi-dialectal tonal language listeners due to coactivation of the lexical representations in both dialects. Tonal information played a significant role in constraining word activation in bi-dialectal auditory word recognition. The results extends our understanding of the role of segment and tone in auditory word recognition in tonal languages from the monolingual context to the bi-dialectal context, and reveals a non-selective processing mechanism in bi-dialect lexical access during auditory word recognition, as in bilingual lexical access.