

**Qian Li**

Institute of Linguistics, Chinese Academy of Social Sciences  
 & Leiden University Center for Linguistics  
[q.li@hum.leidenuniv.nl](mailto:q.li@hum.leidenuniv.nl)

**Yiya Chen**

Leiden University Center for Linguistics  
 & Leiden Institute for Brain and Cognition  
[yiya.chen@hum.leidenuniv.nl](mailto:yiya.chen@hum.leidenuniv.nl)

**Ziyu Xiong**

Institute of Linguistics, Chinese Academy of Social Sciences  
[xiongzuyu@163.com](mailto:xiongzuyu@163.com)

Tianjin Mandarin is a member of the northern Mandarin Chinese family (ISO 693-3: [cmn]). It is spoken in the urban areas of the Tianjin Municipality (CN-12) in the People's Republic of China, which is about 120 kilometers to the southeast of Beijing. Existing studies on Tianjin Mandarin have focused mainly on its tonal aspects, especially its intriguing tone sandhi system, with few studies examining the segmental aspects (on tone, see e.g. Li & Liu 1985, Shi 1986, Liu 1993, Lu 1997, Wang & Jiang 1997, Chen 2000, Liu & Gao 2003, Ma 2005, Ma & Jia 2006, Zhang & Liu 2011, Li & Chen 2016; on segmental aspects, see e.g. Han 1993a, b; Wee, Yan & Chen 2005). As also noted in Wee et al. (2005), this is probably due to the similarity in segmental structures between Tianjin Mandarin and Standard Chinese, especially among speakers of the younger generation, and what differentiates the two Mandarin varieties is most notably their tonal systems. The aim of the present description is therefore to provide a systematic phonetic description of both segmental and tonal aspects of Tianjin Mandarin, with main focus on the tonal aspects.

The sound files illustrated in the present description were produced by a male speaker born in the 1980s. The speaker grew up in the Nankai District of Tianjin, one of the oldest urban districts of Tianjin. He had lived mostly in Tianjin with the exception of a four-year stay in Shanghai for university education. He speaks exclusively Tianjin Mandarin both at home and at work. While the current illustration is based on data from the younger generation, variation between our speaker's generation and older speakers is noted where appropriate. For further details on the generational differences in Tianjin Mandarin, readers are referred to Lu (2004). Note that lexical tones are marked with superscript numbers throughout the paper, i.e. <sup>1</sup> for Tone 1, <sup>2</sup> for Tone 2, <sup>3</sup> for Tone 3 and <sup>4</sup> for Tone 4. (See section 'Lexical tones' for more details on lexical tones in Tianjin Mandarin.)

<sup>1</sup> Syllable structures as in /ŋwei<sup>3</sup>/ and /ɿwei<sup>4</sup>/ cannot be observed in the young-variety Tianjin Mandarin. Sound files for these two syllables are thus not provided.

## Consonants

	Bilabial	Labio-dental	Dento-alveolar	Post-alveolar	Palatal	Alveolo-palatal	Velar
Plosive	p p <sup>h</sup>		t̚ t̚ <sup>h</sup>				k k <sup>h</sup>
Affricate			tʃ̚ tʃ̚ <sup>h</sup>	tʃ̚ tʃ̚ <sup>h</sup>		tʃ̚ tʃ̚ <sup>h</sup>	
Nasal	m		ɳ̚				ŋ
Fricative		f	s̚	ʃ̚		ç	x
Approximant	w			ɻ̚	j ɥ		
Lateral			l̚				

	PHONETIC	ORTHOGRAPHIC	GLOSS
p	p̚e <sup>1</sup>	八	'eight'
p <sup>h</sup>	p̚ <sup>h</sup> e <sup>2</sup>	爬	'to climb'
m	m̚e <sup>1</sup>	妈	'mother'
w	w̚e <sup>1</sup>	蛙	'frog'
f	f̚e <sup>1</sup>	发	'to send'
t̚	t̚e <sup>1</sup>	搭	'to build'
t̚ <sup>h</sup>	t̚ <sup>h</sup> e <sup>1</sup>	他/她	'he/she'
tʃ̚	tʃ̚e <sup>1</sup>	匝	'to circle'
tʃ̚ <sup>h</sup>	tʃ̚ <sup>h</sup> e <sup>1</sup>	擦	'to wipe'
ɳ̚	ɳ̚e <sup>4</sup>	纳	'to include'
s̚	s̚e <sup>1</sup>	撒	'to cast'
l̚	l̚e <sup>1</sup>	拉	'to pull'
tʃ̚	tʃ̚e <sup>1</sup>	渣	'residue'
tʃ̚ <sup>h</sup>	tʃ̚ <sup>h</sup> e <sup>1</sup>	插	'to insert'
ʃ̚	ʃ̚e <sup>1</sup>	沙	'sand'
ɻ̚	ɻ̚ən <sup>2</sup>	人	'person'
tç	tçje <sup>1</sup>	加	'to add'
tç <sup>h</sup>	tç <sup>h</sup> je <sup>1</sup>	掐	'to nip off'
ç	çje <sup>1</sup>	虾	'shrimp'
j	je <sup>1</sup>	鸭	'duck'
ɥ	ɥe <sup>1</sup>	约	'to restrict'
k	k̚ɣ <sup>1</sup>	歌	'song'
k <sup>h</sup>	k̚ <sup>h</sup> ɣ <sup>1</sup>	科	'subject'
ŋ	ɑŋ <sup>2</sup>	昂	'to raise'
x	x̚ɣ <sup>1</sup>	喝	'to drink'

There are 25 consonants in Tianjin Mandarin. To facilitate the comparison of Tianjin Mandarin to Standard Chinese – two closely related Mandarin varieties, we elicited whenever possible the same words as in Lee & Zee (2003) for illustration. Note that although the dentoalveolars are very often not marked with dental diacritics in the literature (e.g. Lee & Zee 2003, Wee et al. 2005), the dental diacritics are marked in the current description following the IPA illustration requirements. (But see Lee & Zee 2003 where a similar 'denti-alveolar' term was used.)

### Plosives

Plosives in Tianjin Mandarin differentiate three places of articulation: bilabial /p p<sup>h</sup>/, as in /p̚e<sup>1</sup>/ 'eight' and /p̚<sup>h</sup>e<sup>2</sup>/ 'to climb', dentoalveolar /t̚ t̚<sup>h</sup>/, as in /t̚e<sup>1</sup>/ 'to build' and /t̚<sup>h</sup>e<sup>1</sup>/ 'he/she',

**Table 1** VOT of unaspirated vs. aspirated plosives in different places of articulation, based on 923 monosyllabic morphemes with plosive onsets.

	Bilabial		Dentoalveolar		Velar	
	Unaspirated	Aspirated	Unaspirated	Aspirated	Unaspirated	Aspirated
Mean	16 ms	91 ms	16 ms	100 ms	36 ms	116 ms
Standard deviation	6 ms	16 ms	4 ms	20 ms	13 ms	22 ms
Number of tokens	121	188	194	145	169	106

and velar /k k<sup>h</sup>/, as in /kɤ<sup>1</sup>/ ‘song’ and /k<sup>h</sup>ɤ<sup>1</sup>/ ‘subject’. They contrast in aspiration, and the contrast holds for all places of articulation. Table 1 shows the mean VOT of aspirated and unaspirated plosives in different places of articulation.

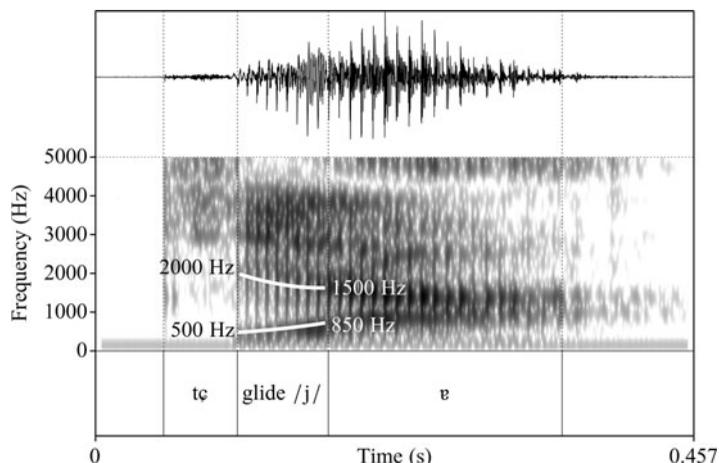
The measurements in Table 1 were made over 923 monosyllabic morphemes with plosive onsets, all of which were selected from a dataset of 3935 monosyllabic morphemes produced by our speaker. Averaged across different places of articulation, the mean VOT for 439 aspirated plosive tokens is 102 ms, while the mean VOT for 484 unaspirated plosive tokens is 23 ms. A one-way ANOVA test was conducted to compare the effect of PLACES OF ARTICULATION (three levels: BILABIAL, DENTOALVEOLAR, and VELAR) on VOT over the 923 plosive tokens. Results revealed a significant effect of PLACES OF ARTICULATION ( $F(2,920) = 20.44, p < .001$ ). A *post-hoc* Tukey HSD test further showed that, velar plosives have significantly longer VOT than bilabial and dentoalveolar plosives (VELAR vs. BILABIAL: Diff. = 20 ms, *p*-adj. < .001; VELAR vs. DENTOALVEOLAR: Diff. = 10 ms, *p*-adj. < .001), while the bilabial and dentoalveolar plosives are not significantly different.

### Affricates

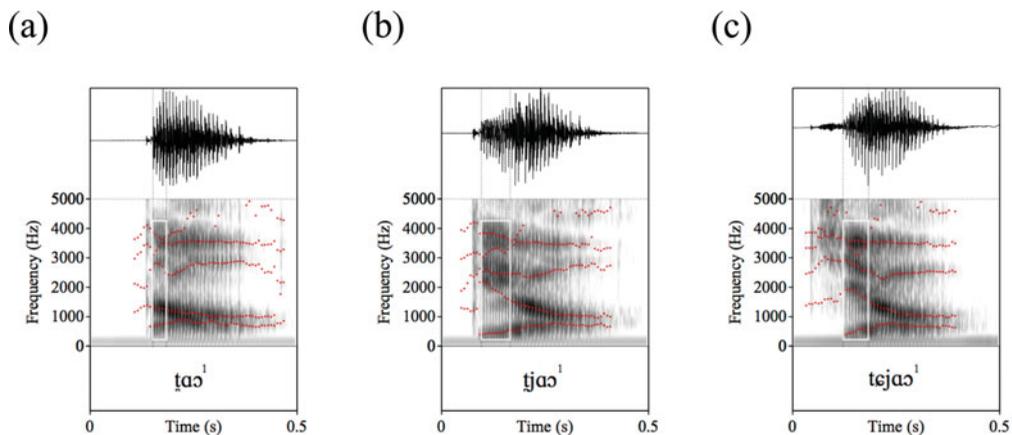
Affricates in Tianjin Mandarin display the same two-way distinction in aspiration as plosives. They have three places of articulation: dentoalveolar /tʂ tʂ<sup>h</sup>/, as in /tʂe<sup>1</sup>/ ‘to circle’ and /tʂ<sup>h</sup>e<sup>1</sup>/ ‘to wipe’, postalveolar /tʃ tʃ<sup>h</sup>/, as in /tʃe<sup>1</sup>/ ‘residue’ and /tʃ<sup>h</sup>e<sup>1</sup>/ ‘to insert’, and alveolo-palatal /tɕ tɕ<sup>h</sup>/ as in /tɕje<sup>1</sup>/ ‘to add’ and /tɕ<sup>h</sup>je<sup>1</sup>/ ‘to nip off’. Like dentoalveolar plosives, the dentoalveolar affricates are produced with the tip of the tongue against the upper front teeth and the tongue blade against the alveolar ridge. The postalveolar affricates are apical, pronounced with the tongue tip raised against the postalveolar region. Both dentoalveolar and postalveolar affricates in Tianjin Mandarin are very similar to that in Standard Chinese (Lee & Zee 2014). The alveolo-palatal affricates are pronounced with the tongue tip down behind the lower front teeth and with the dorsum of the tongue against the area between the alveolar ridge and the hard palate (Ladefoged & Wu 1984).

The postalveolar series are conventionally called ‘retroflexes’ (Chao 1948). However, as discussed in Lee & Zee (2014), this series lacks the action of curling the tongue tip up and back, which is a key feature of typical retroflex articulation (Ladefoged 2006). We thus adopt ‘postalveolar’ rather than ‘retroflex’ for this series. Note that there are some word-specific generational differences in the specific place of articulation for this series of sounds. As observed in Han (1993a, b), while the young-variety Tianjin Mandarin speakers produce some words using the postalveolar consonants (i.e. /tʃ tʃ<sup>h</sup> ʃ/), older-generation speakers typically produce them with their dentoalveolar counterparts (i.e. /tʂ tʂ<sup>h</sup> ʂ/, respectively) (see Wee et al. 2005).

The alveolo-palatal affricates in Tianjin Mandarin are obligatorily followed by a palatal glide if the consonant is not followed by a high vowel, such as in /tɕje<sup>1</sup>/ ‘to add’ and /tɕ<sup>h</sup>je<sup>1</sup>/ ‘to nip off’. Figure 1 shows the spectrogram of /tɕje<sup>1</sup>/ ‘to add’, where we see a transition (glide /j/) between the consonant /tɕ/ and the vowel /e/, taking up between a quarter or a third of the total rhyme length. The F1 of the glide starts with a low value (around 500 Hz) and gradually increases up to about 850 Hz as that of the vowel /e/; the F2 of the glide starts around



**Figure 1** The spectrogram of the whole syllable /tɕjə¹/ 'to add'.



**Figure 2** (Colour online) The spectrograms of /tɕə¹/ 'knife' (a), /tɕjə¹/ 'to hold in the mouth' (b), and /tɕjə¹/ 'to teach' (c). White frames indicate the transitions.

2000 Hz and ends with around 1500 Hz. The F1 and F2 values of the glide onset therefore resemble that of a high front vowel. Glides in such contexts have been traditionally considered part of the rhyme and transcribed as a vowel, such as in /tɕia¹/ and /tɕ<sup>h</sup>ia¹/ (e.g. Wee et al. 2005). In line with Lin (2007), however, we treat /j/ as a glide which constitutes part of the onset.

Figure 2 plots the spectrograms of three monomorphemic words with the same rhyme: /tɕə¹/ 'knife', /tɕjə¹/ 'to hold in the mouth', and /tɕjə¹/ 'to teach'. There is a clear glide-like transition in /tɕjə¹/ (Figure 2(c)), which is similar to that in /tɕjə¹/ in (Figure 2(b)) where there is a real glide. Both are different from that in /tɕə¹/ (Figure 2(a)) where there is only subtle phonetic coarticulation. We take this comparison as additional evidence that in Tianjin Mandarin, there is an underlying glide target between an alveolo-palatal consonant and a non-high vowel. This is different from what is reported by Chen & Gussenhoven (2015) for Shanghai Chinese; there they found only brief phonetic coarticulatory transition exhibited between an alveolo-palatal onset and its following vowel rhyme, suggesting the non-presence of the glide target.

## Nasals

Tianjin Mandarin has nasals in three places of articulation: bilabial /m/, as in /mɛ<sup>1</sup>/ ‘mother’, dentoalveolar /n/, as in /nɛ<sup>4</sup>/ ‘to include’, and velar /ŋ/, as in /ŋɛ<sup>2</sup>/ ‘to raise’. /m/ can only occur in the onset position while /ŋ/ only in the coda position.

## Fricatives

Fricatives in Tianjin Mandarin differentiate five places of articulation: labiodental /f/, as in /fɛ<sup>1</sup>/ ‘to send’, dentoalveolar /s/, as in /sɛ<sup>1</sup>/ ‘to cast’, postalveolar /ʃ/, as in /ʃɛ<sup>1</sup>/ ‘sand’, alveolo-palatal /ç/, as in /çjɛ<sup>1</sup>/ ‘shrimp’, and velar /x/, as in /xɣ<sup>1</sup>/ ‘to drink’. The alveolo-palatal fricative in Tianjin Mandarin is obligatorily followed by a palatal glide (as in /çjɛ<sup>1</sup>/ ‘shrimp’ and /çɥɛ<sup>1</sup>/ ‘boots’) or a high vowel (as in /çi<sup>1</sup>/ ‘west’). (See section ‘Lateral and approximants’ for more details on glides /j, ɥ/.) The velar fricative /x/ is realized with the uvular fricative [χ] when followed by a low vowel, as in /xæ<sup>2</sup>/ ([χæ<sup>2</sup>]) ‘child’ and /xɑɔ<sup>3</sup>/ ([χɑɔ<sup>3</sup>]) ‘good’.

## Lateral and approximants

Tianjin Mandarin has one lateral /l/, as in /lɛ<sup>1</sup>/ ‘to pull’, and four approximants: /w ɹ j ɥ/, as in /wɛ<sup>1</sup>/ ‘frog’, /ɹɛ<sup>2</sup>/ ‘person’, /jɛ<sup>1</sup>/ ‘duck’, and /ɥɛ<sup>1</sup>/ ‘to restrict’. Among the four approximants, /w j ɥ/ can serve as syllable onset or part of a complex onset. As a syllable onset, /w/ is sometimes pronounced as the labiodental voiced consonant [v] as also reported in Han (1993b). /j/ and /ɥ/ are both palatal, with the main contrast in lip rounding. They do not occur in the same contexts except before the vowel /e/, as in /njɛ<sup>1</sup>/ ‘to pinch’ vs. /nɥɛ<sup>4</sup>/ ‘to abuse’, /tçjɛ<sup>1</sup>/ ‘to connect’ vs. /tçɥɛ<sup>2</sup>/ ‘to feel’, /tç<sup>h</sup>jɛ<sup>1</sup>/ ‘to cut’ vs. /tç<sup>h</sup>ɥɛ<sup>1</sup>/ ‘to lack’, and /çjɛ<sup>1</sup>/ ‘to rest’ vs. /çɥɛ<sup>1</sup>/ ‘boots’. According to Han (1993a, b), as also noted by one of the reviewers, the onset /ɹ/ can be pronounced as [j] in the old variety of Tianjin Mandarin in words such as /ɹəʊ<sup>4</sup>/ ([jəʊ<sup>4</sup>]) ‘meat’; and interestingly, the onset /j/ can also be pronounced as [ɹ] in words such as /jɑŋ<sup>4</sup>/ ([ɹɑŋ<sup>4</sup>]) ‘to brim over’. Further research is needed to understand the lexical-specific swap. Furthermore, /ɥ/ was reported to be produced as [j] in the old variety (Han 1993b).

## Syllabic consonants

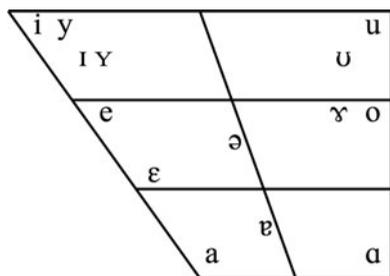
Tianjin Mandarin has two syllabic consonants (notation in line with Chao 1948): the dentoalveolar /ɹ̥/, as in /tɕɹ̥<sup>3</sup>/ ‘son’, and postalveolar /ʃ̥/, as in /tʃ̥<sup>3</sup>/ ‘paper’. These two phonemes are traditionally referred to as ‘apical vowels’ by Sinologists (e.g. Karlgren 1915–1926). Ladefoged & Maddieson (1996) refer to them as ‘fricative vowels’. With evidence from both ultrasound imaging and acoustic data, Lee-Kim (2014) shows that these two phonemes are neither vowels nor fricative, but more comparable with approximants in nature, i.e. syllabic dental approximant /ɹ̥/ and retroflex approximant /ʃ̥/, respectively. This is similar to observations made in Lee & Zee (2003, 2014), in which, however, the two sounds have been transcribed with the same syllabic approximant /ɹ̥/.

The dentoalveolar /ɹ̥/ only follows dentoalveolar consonants /tɕ tɕ<sup>h</sup> s/ (as in /tɕɹ̥<sup>3</sup>/ ‘son’, /tɕ<sup>h</sup>ɹ̥<sup>2</sup>/ ‘word’, /sɹ̥<sup>1</sup>/ ‘to think’), while /ʃ̥/ follows postalveolar consonants /tʃ tʃ<sup>h</sup> ʃ/ (as in /tʃ̥<sup>3</sup>/ ‘paper’, /tʃ<sup>h</sup>ɹ̥<sup>1</sup>/ ‘to eat’, /ʃ̥ɹ̥<sup>1</sup>/ ‘poem’). Their tongue configurations are similar to the preceding homorganic consonants, i.e. /ɹ̥/ is homorganic with /tɕ, tɕ<sup>h</sup>, s/, and /ʃ̥/ is homorganic with /tʃ tʃ<sup>h</sup> ʃ/. In addition, /ɹ̥/ can occur by itself as in /ɹ̥<sup>4</sup>/ ‘the sun’. To highlight the homorganicity of the preceding consonants and the following syllabic approximants, we transcribe the two syllabic consonants with two independent symbols as in Lee-Kim (2014), but we adopted the postalveolar /ʃ̥/ symbol rather than the retroflex /ɹ̥/.

For more information on the debates regarding both the phonemic status and notation of the two sounds, as well as their acoustic and articulatory realizations, readers are referred to Lee-Kim (2014) and Lee & Zee (2014).

## Vowels

### Monophthongs

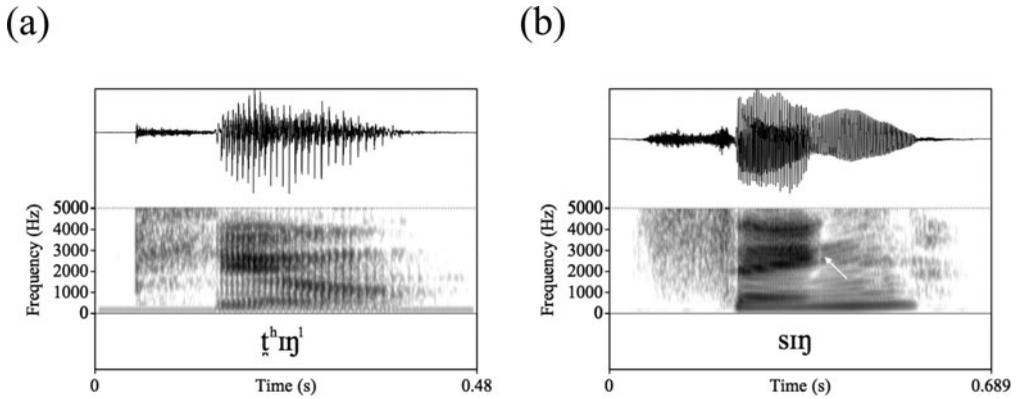


	PHONETIC	ORTHOGRAPHIC	GLOSS
i	tɿ <sup>1</sup>	低	'low'
y	ly <sup>2</sup>	驴	'donkey'
e	tje <sup>1</sup>	爹	'dad'
ɐ	tɐ <sup>1</sup>	搭	'to build'
o	tʷo <sup>1</sup>	多	'more'
ɤ	tɤ <sup>2</sup>	得	'to get'
u	tu <sup>1</sup>	都	'metropolis'
ə	tə	的	Possessive marker
	tən <sup>4</sup>	拖	'to drag'
	tən <sup>1</sup>	灯	'lamp'
ɪ	lɪn <sup>2</sup>	林	'forest'
	lɪŋ <sup>2</sup>	零	'zero'
ʏ	tɕʏn <sup>1</sup>	军	'army'
ɛ	tjɛn <sup>1</sup>	颠	'bump'
a	ʈan <sup>1</sup>	单	'single'
ɑ	ʈan <sup>1</sup>	当	'when'
ʊ	ʈɔŋ <sup>1</sup>	东	'east'

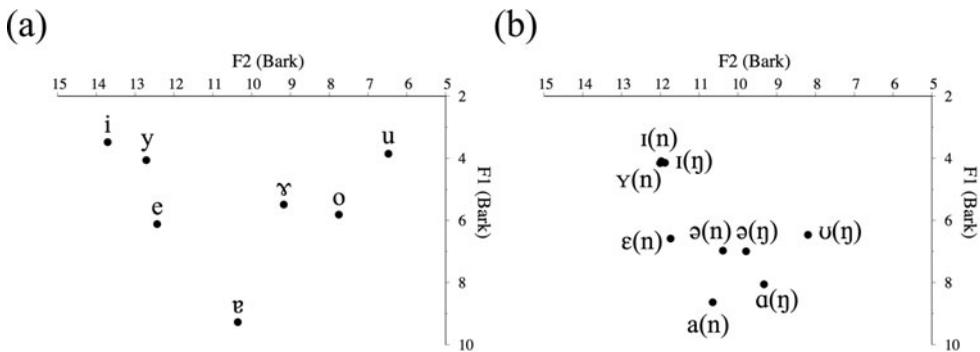
Tianjin Mandarin has 14 monophthongs. /i y e ɐ ɤ u/ occur in open syllables, /ɪ ʏ ɛ ɑ ʊ/ occur in closed syllables, and /ə/ occurs in both open and closed syllables.

Among the seven vowels that occur only in open syllables, /i/ and /y/ (as in /tɿ<sup>1</sup>/ 'low' and /ly<sup>2</sup>/ 'donkey') are high front vowels contrasting in lip rounding. /e/ is a mid-high front vowel which obligatorily follows an onset glide (i.e. /j ɥ w/), as in /tje<sup>1</sup>/ 'dad'. /ɐ/ is a low mid vowel, as in /tɐ<sup>1</sup>/ 'to build'. /u/ is a high back rounded vowel, as in /tu<sup>1</sup>/ 'metropolis'. Both /ɤ/ and /o/ are mid-high back vowels differing mainly in lip rounding, as in /tɤ<sup>2</sup>/ 'to get' and /tʷo<sup>1</sup>/ 'more' (where an onglide /w/ is obligatory before /o/). (But see Han 1993b, which reported that all syllables with /o/ are pronounced with /ɤ/ by speakers of the old variety of Tianjin Mandarin.)

/ɪ/ (as in /lɪn<sup>2</sup>/ 'forest' and /lɪŋ<sup>2</sup>/ 'zero') and /ʏ/ (as in /tɕʏn<sup>1</sup>/ 'army') both occur only in closed syllables. They are the lax counterparts of /i/ and /y/, respectively. /ɪ/ occurs before both dentoalveolar and velar nasal coda, while /ʏ/ only occurs before dentoalveolar nasal coda. When /ɪ/ and /ʏ/ are followed by nasal codas, an offglide [ə] is inserted as the articulation of the vowel transits to the nasal in the following coda. This is illustrated in Figure 3, where the spectrogram of the syllable /t<sup>h</sup>ɪŋ<sup>1</sup>/ ([t<sup>h</sup>ɪŋ<sup>1</sup>] 'to listen') in Tianjin Mandarin is plotted against that of the syllable /sɪŋ/ 'sing' in American English (Ladefoged 1999). In the latter, there is a clear and sharp acoustic boundary between the vowel and the nasal coda



**Figure 3** Spectrogram comparison of the syllable /tʰɪŋ¹/ in Tianjin Mandarin (a) vs. the syllable /sɪŋ/ in American English (b).



**Figure 4** F1 and F2 values (in Bark) of monophthongs produced in open syllables (a) and in closed syllables (b).

(as shown with an arrow in [Figure 3\(b\)](#)) without the presence of a transitional schwa. One of our reviewers pointed out that the nasal codas in Tianjin Mandarin are possibly nasal glides while the English nasal codas are nasal stops (see also Wang 1997), which – we agree – might be the reason behind the vowel–coda juncture difference. Further studies are needed to verify this possibility.

/ɛ/ is a mid-low front vowel which only occurs before the dentoalveolar nasal coda, as in /tʃɛŋ¹/ ‘bump’. /ɛ/ is often treated as an allophone of /a/ in traditional descriptions of Tianjin Mandarin, for example in Han (1993a) and Wee et al. (2005). /a/ and /ɑ/ are both low vowels, which, however, occur in different contexts. The low front /a/ occurs before dentoalveolar nasal coda (as in /tʃaŋ¹/ ‘single’), and the low back vowel /ɑ/ before the velar nasal coda (as in /tʃɑŋ¹/ ‘when’). /ʊ/ is the lax counterpart of /u/, which occurs before the velar nasal coda as in /tʃʊŋ¹/ ‘east’.

/ə/ is a central vowel which can appear in both open and closed syllables although an open syllable with /ə/ is exclusively a neutral-tone syllable (e.g. /tə/ POSSESSIVE MARKER in /wo³ tə/ ‘mine’). (See section ‘Neutral tone’ below for more details on neutral tone.) In closed syllables, /ə/ can occur before both nasal codas /ŋ/ and /ŋ/ (as in /təŋ⁴/ ‘to drag’ and /təŋ¹/ ‘lamp’). Before the dentoalveolar nasal coda, /ə/ is slightly more fronted than before the velar nasal coda /ŋ/.

[Figure 4](#) shows the mean F1 and F2 values of each monophthong occurring in open and closed syllables. The formant data in [Figure 4\(a\)](#) were based on 50 samples of each monophthong produced in open syllables by measuring the vowel midpoint. /ə/ is not included

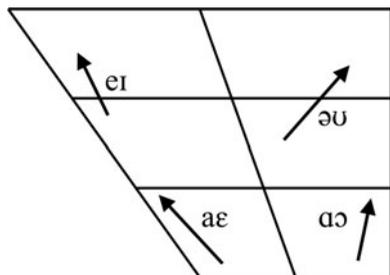
in Figure 4(a) because open syllables with /ə/ can only occur in neutral-tone syllables, which do not occur in isolation. Monophthongs occurring in closed syllables with dentoalveolar and velar nasal coda are plotted in Figure 4(b). Formant values in both graphs were converted from Hertz to Bark using the following formula (after Boersma & Weenink 2017):

$$Bark = 7 \times \ln \left( \frac{Hz}{650} + \sqrt{\left( \frac{Hz}{650} \right)^2 + 1} \right)$$

It can be seen from Figure 4 that vowels in closed syllables (Figure 4(b)) are more centralized compared to those in open syllables (Figure 4(a)). Furthermore, the realization of central vowel /ə/ is influenced by different following nasal codas due to their closure gestures at different places of articulation. To be specific, /ə/ is more fronted if followed by the dentoalveolar nasal coda /ŋ/, but more backward if followed by the velar nasal coda /ŋ/.

Note that vowels in closed syllables have been often treated as allophonic variants of vowels in open syllables, mainly based on the fact that they are mutually non-contrastive, and occur in different contexts. For example, Wee et al. (2005) use /a/ for both open syllables (e.g. /p<sup>h</sup>a<sup>2</sup>/ ‘to climb’) and closed syllables (e.g. /san<sup>1</sup>/ ‘three’ and /taŋ<sup>3</sup>/ ‘party’). However, given the clearly different vowel quality, here we adopt an alternative view to treat them as different phonemes (also see Lin 2007 and Chen & Gussenhoven 2015 for similar treatments). This is to highlight the phonological non-equivalence of pairs of open-syllable vowel vs. closed-syllable vowel (Chen & Gussenhoven 2015).

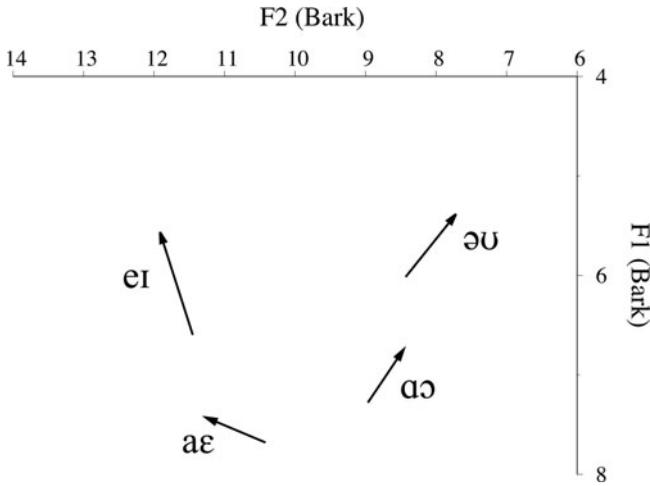
### Diphthongs



	PHONETIC	ORTHOGRAPHIC	GLOSS
ei	ɿei <sup>2</sup>	雷	‘thunder’
aε	ʈaε <sup>1</sup>	呆	‘dull’
ɑɔ	ʈɑɔ <sup>1</sup>	刀	‘knife’
əu	ʈəu <sup>1</sup>	都	‘all’

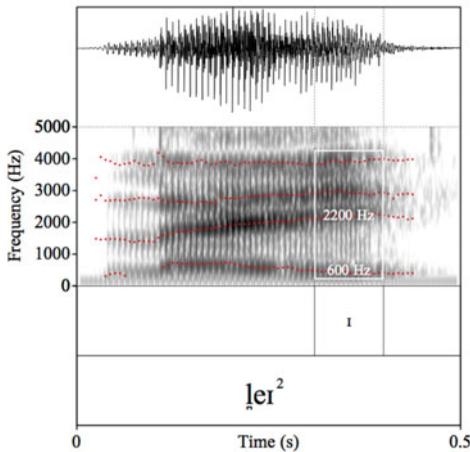
There are four diphthongs in Tianjin Mandarin, with /ei aε/ gliding towards the front (as in /ɿei<sup>2</sup>/ ‘thunder’ and /ʈaε<sup>1</sup>/ ‘dull’) and /ɑɔ əu/ towards the back (as in /ʈɑɔ<sup>1</sup>/ ‘knife’ and /ʈəu<sup>1</sup>/ ‘all’). All diphthongs only occur in open syllables. Figure 5 shows the mean F1 and F2 values of 50 samples for each diphthong by measuring the respective midpoint of the two parts in the vowel. All samples were selected from the 3935 monosyllabic words produced by our speaker. Arrows in Figure 5 demonstrate the trajectories of the gliding. Formant data were converted from Hertz to Bark using the following formula (after Boersma & Weenink 2017):

$$Bark = 7 \times \ln \left( \frac{Hz}{650} + \sqrt{\left( \frac{Hz}{650} \right)^2 + 1} \right)$$

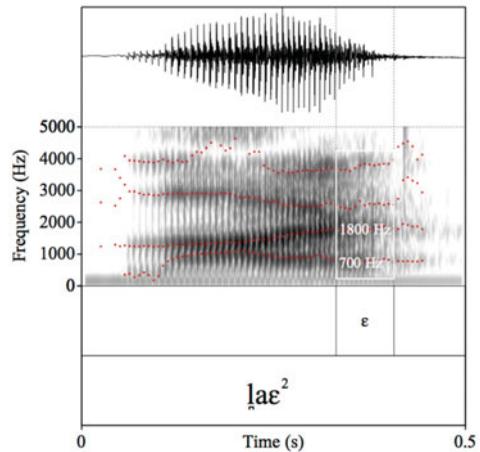


**Figure 5** F1 and F2 values (in Bark) of diphthongs. Gliding trajectories are shown with arrows.

(a)



(b)

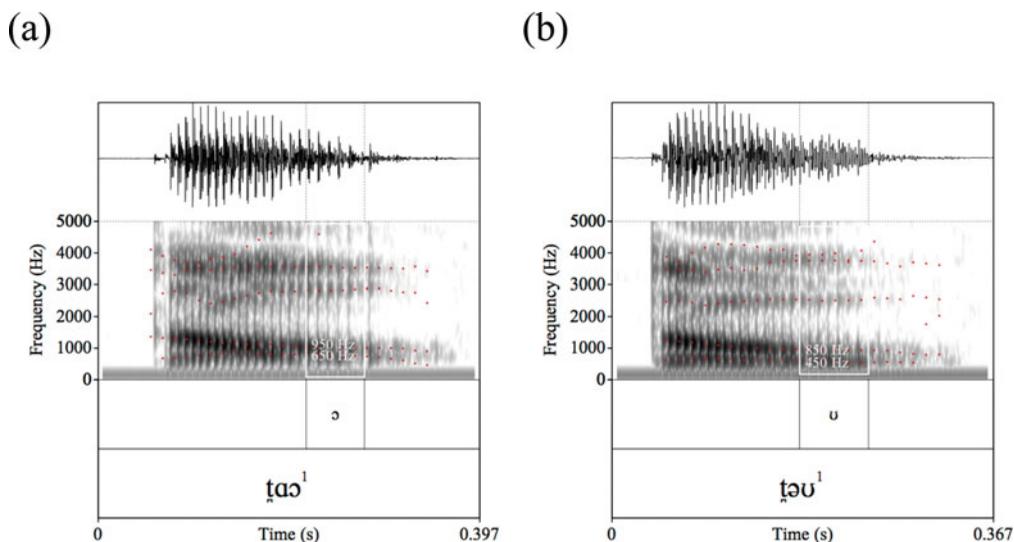


**Figure 6** (Colour online) Spectrogram comparison of diphthongs /eɪ/ vs. /aɛ/ as in syllables /ɬeɪ²/ 'thunder' (a) vs. /ɬaɛ²/ 'to come' (b), respectively.

/eɪ/ and /aɛ/ are frequently transcribed as /ei/ and /ai/, respectively, both gliding towards the same high front target /i/ (e.g. Han 1993a, b; Wee et al. 2005). Figure 5 shows that neither /eɪ/ nor /aɛ/ in Tianjin Mandarin really reaches the high front region at the offset part.

To further illustrate the different end points of /eɪ/ and /aɛ/, Figure 6 compares their spectrograms. As shown in Figure 6, the end points of /eɪ/ vs. /aɛ/ are very different especially in terms of F1 and F2, where the offset part of /eɪ/ shows clearly lower F1 (about 600 Hz) but higher F2 values (about 2200 Hz) than that of /aɛ/ (F1: about 700 Hz; F2: about 1800 Hz).

Similar differences can also be observed for /aɔ/ vs. /əʊ/, both of which have been frequently described as gliding towards the high back vowel /u/ (e.g. Han 1993a, b; Wee et al. 2005). Figure 7 compares the spectrograms of /aɔ/ vs. /əʊ/, which again show different



**Figure 7** (Colour online) Spectrogram comparison of diphthongs /aɔ/ vs. /əʊ/ as in syllables /tʰaɔ¹/ 'knife' (a) vs. /tʰəʊ¹/ 'all' (b), respectively.

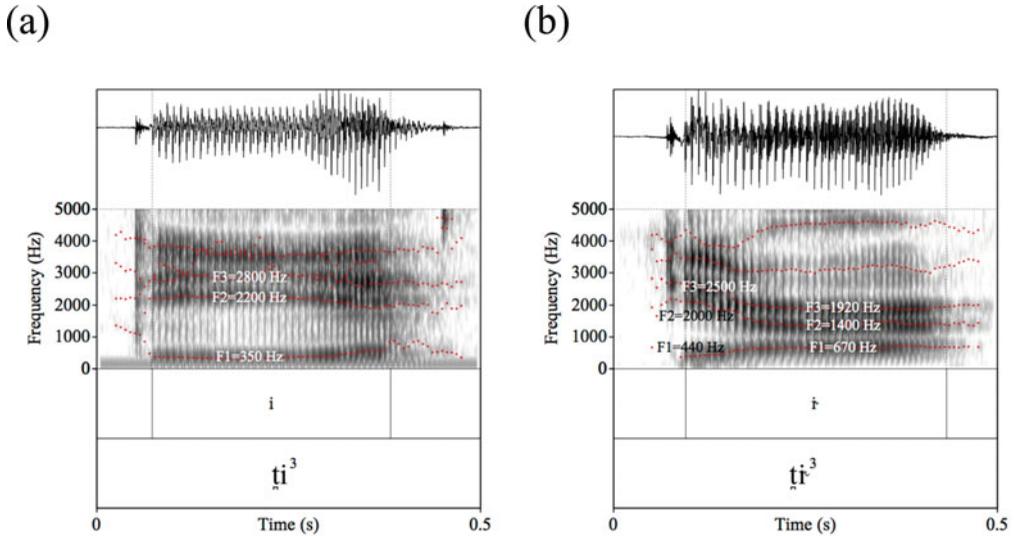
qualities of the two vowels at the end. To be specific, /əʊ/ (Figure 7(b)) shows lower F1 (450 Hz) and F2 (850 Hz) than those of /aɔ/ (F1: 650 Hz; F2: 950 Hz; see Figure 7(a)).

### Rhotic vowel and er-hua

Tianjin Mandarin has a rhotic vowel /əʀ/, which is produced as an r-colored schwa with the tip of tongue raised. /əʀ/ is syllabic, as in /əʀ²/ 'son', /əʀ³/ 'ear' and /əʀ⁴/ 'two'. When /əʀ/ is produced in a neutral-tone syllable, it is used as a diminutive suffix.

When adding the suffix /əʀ/ to a noun, the two syllables are typically coalesced into one rhoticized syllable in the output form. The vowel part of the preceding syllable is directly rhoticized and only the lexical tone of the preceding syllable is kept, e.g. /tʰi³/ + /əʀ/ → [tʰi³] 'remnants' (as compared to [tʰi³] 'bottom'). Such process is known as 'rhoticization' or 'er-hua' in Chinese. Figure 8 compares the spectrograms of non-rhoticized [i] to its rhoticized counterpart [iʀ]. As shown in Figure 8(b), the first quarter of [iʀ] is realized similarly to [i] (as in Figure 8(a)). The remaining three-quarters, however, is realized with a clearly lowered mean F3 (from about 2500 Hz to 1920 Hz), which is a typical acoustic cue of rhoticization (Ladefoged & Maddieson 1996). In addition, the process of rhoticization often changes the F1 and F2 values. For example, in Figure 8(b), the F1 of /i/ has been changed from 440 Hz to 670 Hz and F2 from 2000 Hz to 1400 Hz. Due to the changes in F1 and F2, rhoticization has been traditionally transcribed as /əʀ/-insertion, e.g. /i/ + /əʀ/ → [iəʀ] (see e.g. Han 1993a). However, given the clear F3 lowering in the rhoticized vowel, we regard the F1 and F2 changes as the by-products of rhoticization rather than a rhoticized schwa insertion.

There are also other rhoticizing processes in both open and closed syllable structures. In open syllables, if the vowel is /ae/, only the /a/ part is rhoticized while /e/ is deleted, e.g. /pʰae²/ + /əʀ²/ → [pʰa²] 'badge' (as compared to [pʰae²] 'card'); if the rhyme is /e/, /ei/ or a syllabic consonant, the entire rhyme part is replaced with /əʀ/, e.g. /per⁴/ + /əʀ²/ → [pəʀ⁴] 'very' (as compared to [pe⁴] 'double'). In closed syllables, if the coda is /ŋ/, the vowel is rhoticized, while the coda is deleted, e.g. /paŋ⁴/ + /əʀ²/ → [pəʀ⁴] 'partner' (as compared to [paŋ⁴] 'companion'); if the coda is /ŋ/, the vowel is nasalized and rhoticized, e.g. /xwəŋ²/ + /əʀ²/ → [xwəʀ²] 'yolk' (as compared to [xwəŋ²] 'yellow').



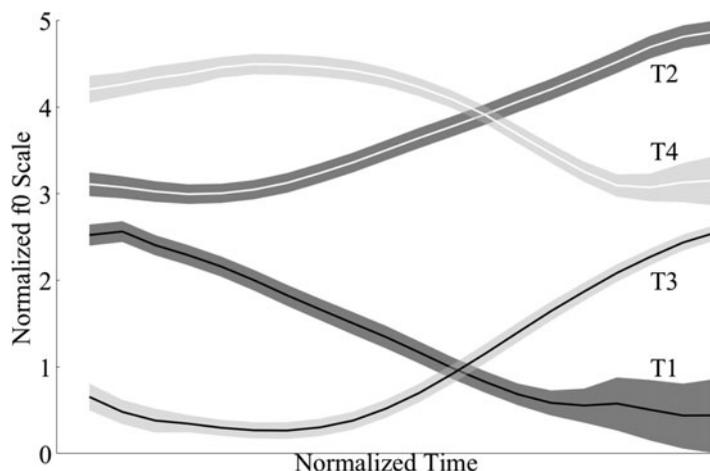
**Figure 8** (Colour online) Spectrogram comparison of non-rhotacized [i] vs. rhotacized [ɿ] as in syllables [ti<sup>3</sup>] ‘bottom’ (a) vs. [tɿ<sup>3</sup>] ‘remnants’ (b), respectively.

### Syllable structure and phonotactics

The syllable structure in Tianjin Mandarin is (C)(G)V(C). C stands for consonant, G for glide, and V for vowel. Except for /ŋ/, all consonants can occur at syllable onset. Glides can also serve as syllable onset, such as /j/ in /jɐ<sup>1</sup>/ ‘duck’, /w/ in /wɐ<sup>1</sup>/ ‘frog’ and /ɥ/ in /ɥɛ<sup>1</sup>/ ‘to restrict’. Onsetless syllables are also possible, as in /æ<sup>1</sup>/ ‘sad’. However, in the old variety Tianjin Mandarin as reported in Han (1993b), onsetless syllables are not allowed before the rhymes of /aŋ/, /aŋ/, /əŋ/, /ɤ/, /æ/, /ɑ/, or /əu/; a nasal onset /ŋ/ is obligatory. Coda is optional, and only /ŋ/ and /ŋ/ are allowed (as in /tʰaŋ<sup>1</sup>/ ‘single’ or /tʰaŋ<sup>1</sup>/ ‘when’). /ŋ/ is the only consonant in the language that can appear both at the beginning and end of a syllable (as in /ŋɐ<sup>4</sup>/ ‘to include’ and /tʰaŋ<sup>1</sup>/ ‘single’).

In addition, there are some co-occurrence restrictions of consonants and vowels in Tianjin Mandarin:

- (i) High front vowels /i y ɪ ɤ/ and the corresponding glides /j ɥ/ cannot follow dentoalveolar obstruents /tʂ tʂ<sup>h</sup> ʂ/, postalveolar /tʃ tʃ<sup>h</sup> ʃ/, or velar consonants /k k<sup>h</sup> x/. However, /y/ and /ɥ/ can follow dentoalveolar /ɲ l/, as in /ɲy<sup>3</sup>/ ‘female’, /lɥ<sup>2</sup>/ ‘donkey’, /ɲɥɛ<sup>4</sup>/ ‘to abuse’, /lɥɛ<sup>4</sup>/ ‘to omit’.
- (ii) Alveolo-palatals /tɕ tɕ<sup>h</sup> ɕ/ can only occur before high front vowels /i y ɪ ɤ/, as in /tɕi<sup>1</sup>/ ‘chicken’, /tɕ<sup>h</sup>i<sup>1</sup>/ ‘seven’, /ɕi<sup>1</sup>/ ‘west’, /tɕy<sup>1</sup>/ ‘to live’, /tɕ<sup>h</sup>y<sup>1</sup>/ ‘maggot’, /ɕy<sup>1</sup>/ ‘needs’, /tɕɪŋ<sup>1</sup>/ ‘gold’, /tɕ<sup>h</sup>ɪŋ<sup>1</sup>/ ‘to invade’, /ɕɪŋ<sup>1</sup>/ ‘heart’, /tɕɤŋ<sup>1</sup>/ ‘army’, /tɕ<sup>h</sup>ɤŋ<sup>2</sup>/ ‘skirt’, /ɕɤŋ<sup>1</sup>/ ‘to fumigate’, or their corresponding glides /j ɥ/, as in /tɕje<sup>1</sup>/ ‘to connect’, /tɕ<sup>h</sup>je<sup>1</sup>/ ‘to cut’, /ɕje<sup>1</sup>/ ‘to rest’, /tɕɥɛ<sup>2</sup>/ ‘to feel’, /tɕ<sup>h</sup>ɥɛ<sup>1</sup>/ ‘to lack’, /ɕɥɛ<sup>1</sup>/ ‘boots’. Given this restriction, the phonemic status of the alveolo-palatals has been a matter of debate, as relative to the dentoalveolar /tʂ tʂ<sup>h</sup> ʂ/, the postalveolar /tʃ tʃ<sup>h</sup> ʃ/, and the velar /k k<sup>h</sup> x/. A full discussion of their status is beyond the scope of the present paper; interested readers are referred to Lin (2014) for further details.
- (iii) Mid-high vowels /e o/ and mid-low vowel /ɛ/ have to co-occur with glides /j w/, respectively, as in /tje<sup>1</sup>/ ‘dad’, /tʰwo<sup>1</sup>/ ‘more’, /tjeŋ<sup>1</sup>/ ‘bump’. Mid-high back vowel /ɤ/ cannot follow labial consonants /p p<sup>h</sup> m f/ in the younger-variety Tianjin Mandarin; the structure of /p p<sup>h</sup> m f/ + /ɤ/ has been only reported for the old variety in Han (1993b).



**Figure 9** Lexical tones in isolation. Lines stand for the mean. Gray areas stand for  $\pm 1$  standard error of mean. Tone 1 (T1) is illustrated with black line and dark gray area; Tone 2 (T2) with white line and dark gray area; Tone 3 (T3) with black line and light gray area; Tone 4 (T4) with white line and light gray area. Time is normalized.

- (iv) Glide /w/ cannot be followed by front vowels while /ɥ/ can only precede front vowels, as in /ɥe<sup>1</sup>/ ‘to restrict’. /j/, however, can be followed by front, central and back vowels, as in /ji<sup>1</sup>/ ‘one’, /jɐ<sup>1</sup>/ ‘duck’, /jɑŋ<sup>1</sup>/ ‘central’.
- (v) Han (1993b) noted that in the old-variety Tianjin Mandarin, /ŋ/ or /l/ do not co-occur with the rounded high front vowel /y/ in a number of open syllables. So, words like /ŋy<sup>3</sup>/ ‘female’ and /ly<sup>4</sup>/ ‘green’ in the young-variety Tianjin Mandarin would be pronounced as /ŋwer<sup>3</sup>/ and /lwei<sup>4</sup>/ in the old-variety Tianjin Mandarin.<sup>1</sup>

## Tones

### Lexical tones

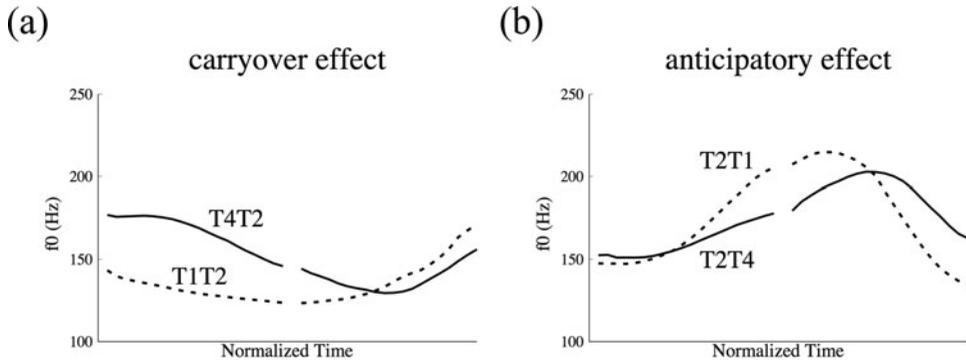
	PHONETIC	ORTHOGRAPHIC	GLOSS
Tone 1	lɑɔ <sup>1</sup>	捞	‘to dredge up’
Tone 2	lɑɔ <sup>2</sup>	劳	‘hard-working’
Tone 3	lɑɔ <sup>3</sup>	老	‘old’
Tone 4	lɑɔ <sup>4</sup>	涝	‘to flood’

There are four full lexical tones in Tianjin Mandarin. Figure 9 shows the f<sub>0</sub> contours of the four lexical tones elicited in isolation with obstruent onsets. Each tonal contour was obtained by averaging across 50 samples, all of which were selected from the 3935 monosyllabic words produced by our speaker. The f<sub>0</sub> values were normalized so that f<sub>0</sub> can be interpreted into the five-scale pitch system using the T-normalization method developed by Shi (1986). The intervals 0–1, 1–2, 2–3, 3–4, and 4–5 correspond to pitch levels 1–5 in Chao’s (1920) lexical tone annotation system, respectively.

As illustrated in Figure 9, Tone 1 (hereafter referred to as T1) is a low-falling tone, of which pitch contour falls from the mid to the lower end of the speaker’s pitch range, as in /lɑɔ<sup>1</sup>/ ‘to dredge up’. Tone 2 (T2) is a high-rising tone, whose pitch contour rises from the mid to the upper end of the pitch range, as in /lɑɔ<sup>2</sup>/ ‘hard-working’. Tone 3 (T3) is a low-dipping or low-rising tone, which falls slightly from the lower pitch range, stays at the bottom and then rises to the mid pitch range of the speaker, as in /lɑɔ<sup>3</sup>/ ‘old’. Tone 4 (T4) is a high-falling

**Table 2** Transcriptions of the four lexical tones in Tianjin Mandarin in different studies (Li & Chen 2016).

T1	T2	T3	T4	Studies
L	H	R(LH)	F(HL)	e.g. Chen 2000, Wee et al. 2005, Hyman 2007
LL	HH	LH	HL	e.g. Wang & Jiang 1997, Wang 2002, Ma 2005
21	45	213	53	e.g. Li & Liu 1985, Hung 1987, Tan 1987, Zhang 1987, Han 1993a
211	455	113	553	e.g. Shi 1986
41	34	12	52	e.g. Zhang & Liu 2011

**Figure 10**  $f_0$  realization of T2 when connected with a high tone (T4) vs. a low tone (T1). T2 as the second tone in (a), T2 as the first tone in (b). Time is normalized.

tone which falls from the upper end to the mid of the pitch range, as in / $\text{la}\text{a}^4$ / ‘to flood’. It is noticeable that T1 and T4 differ in the overall tonal height where T1 is realized in a lower pitch range while T4 in a relatively higher one. Furthermore, T4 has a high plateau/rise at the beginning, while T1 does not. Adopting the pitch range scale in Chao (1920), T1 can be transcribed as /31/, T2 as /45/, T3 as /213/ or /13/, T4 as /53/.

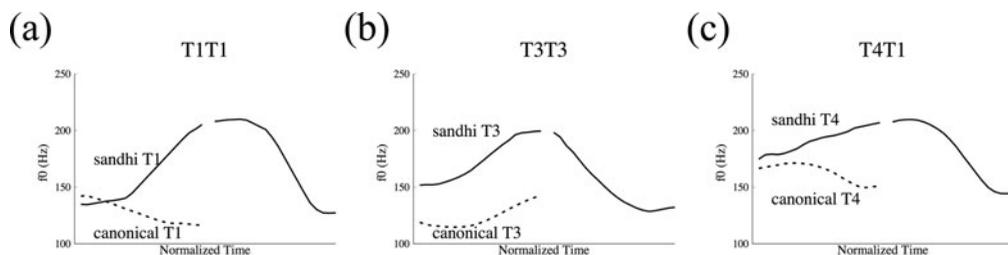
Previous studies on Tianjin lexical tones have been mainly based on impressionistic observations. Researchers have varied greatly in their annotation of the four lexical tones, as summarized in Table 2. It is worth noting that although most studies differ in the absolute pitch values for the four tones, at a more abstract level, the basic  $f_0$  patterns of the four lexical tones in Tianjin Mandarin can be described as low-falling (T1), high-rising (T2), low-dipping/low-rising (T3), and high-falling (T4).

### Tonal variability

When lexical tones are produced in connected speech, their  $f_0$  realizations usually deviate from the canonical tonal contours that are produced in isolation, due to different contextual tonal variation processes such as tonal coarticulation and tone sandhi.

### Tonal coarticulation

Tonal coarticulation in Tianjin Mandarin is bi-directional, including the left-to-right carryover effects as well as the right-to-left anticipatory effects. Carryover tonal coarticulation in Tianjin Mandarin is assimilatory in nature, while anticipatory tonal coarticulation tends to be dissimilatory. In Tianjin Mandarin, the carryover tonal coarticulation can be observed in all tonal contexts except when the second tone is the low-falling T1 (Zhang & Liu 2011), while the anticipatory tonal coarticulation is only triggered by low tones (i.e. T1 and T3) (Li & Chen 2016; but see Zhang & Liu 2011, which reports anticipatory effect only before T3). Figure 10 illustrates the two coarticulatory effects in Tianjin Mandarin. Each tonal contour was obtained by averaging across 12 disyllabic samples produced by our speaker. For more



**Figure 11**  $f_0$  realization of three tone sandhi patterns in Tianjin Mandarin. Solid lines for the  $f_0$  realization of the sandhi patterns, dashed lines for the canonical  $f_0$  realization. Time is normalized.

details on tonal coarticulation in Tianjin Mandarin, readers are referred to Zhang & Liu (2011) and Li & Chen (2016).

As shown in Figure 10(a), the  $f_0$  of a tone can be realized differently due to different preceding tones: when T2 is preceded by a high tone such as T4 (as in /tɕi<sup>4</sup> mɔu<sup>2</sup>/ ‘stratagem’), the onset  $f_0$  realization of the second T2 is clearly higher than that following a low tone such as T1 (as in /kwɛi<sup>1</sup> mwo<sup>2</sup>/ ‘scale’). Similar carryover effects could be observed from the comparison of T4T4 (as in /ŋɛi<sup>4</sup> mu<sup>4</sup>/ ‘inside story’) vs. T1T4 (as in /kwɛi<sup>1</sup> mi<sup>4</sup>/ ‘best female friend’).

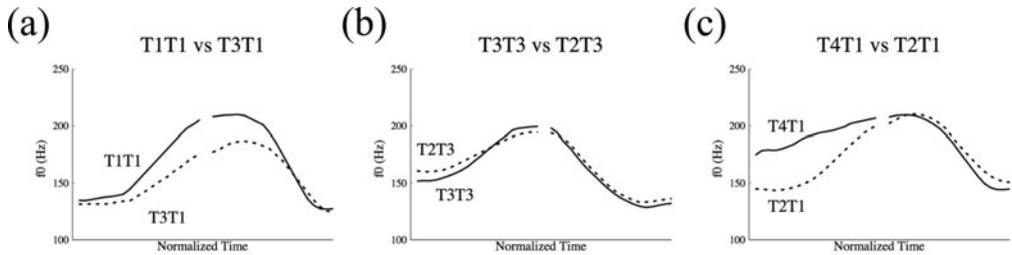
Figure 10(b) illustrates the anticipatory effects, where the first tone is realized differently due to different following tones: when T2 is followed by a low tone such as T1 (as in /pæi<sup>2</sup> mɔu<sup>1</sup>/ ‘white cat’), the offset  $f_0$  realization of the first T2 shows faster rate of  $f_0$  rise than that before a high tone such as T4 (as in /tʃ<sup>h</sup>əu<sup>2</sup> mi<sup>4</sup>/ ‘dense’). Similar anticipatory effects could also be observed from the comparison of T4T2 (as in /tɕi<sup>4</sup> mɔu<sup>2</sup>/ ‘stratagem’) vs. T4T3 (as in /mi<sup>4</sup> mɛ<sup>3</sup>/ ‘password’).

### *Tone sandhi*

Previous impressionistic studies on Tianjin Mandarin have proposed four disyllabic tone sandhi patterns: T1T1, T3T3, T4T1, and T4T4 (e.g. Li & Liu 1985, Hung 1987, Tan 1987, Zhang 1987, Chen 2000, Wang 2002, Hyman 2007, but see Wee et al. 2005 for two more disyllabic sandhi patterns: T3T2 and T3T4). Among the four claimed tone sandhi patterns, only three have been confirmed with experimental data (Li & Chen 2016): T1T1, T3T3, T4T1. Figure 11 shows the  $f_0$  contours of the three tone sandhi patterns, with each tonal contour obtained by averaging across 12 samples produced by our speaker. These patterns have been further verified by an experimental study of five more speakers in Li & Chen (2016).

It can be seen from Figure 11 that in all the three tonal combinations, the first tone is realized with a drastically different  $f_0$  contour from that of its canonical form (compared to their respective contours in Figure 9 plotted here as the dashed lines). In T1T1, the first T1 does not have a low-falling  $f_0$  contour any more as its canonical form (dashed line in Figure 11(a)). Instead, the  $f_0$  offset of the sandhi tone (solid line in Figure 11(a)) is raised to a great extent (as in /tɕjɛ<sup>1</sup> mɔu<sup>1</sup>/ ‘domestic cat’). In T3T3, the first T3 is realized with a high-rising  $f_0$  contour (solid line in Figure 11(b)), which is different from the low-dipping/low-rising  $f_0$  contour of the canonical T3 (dashed line in Figure 11(b)), as in /wu<sup>3</sup> ny<sup>3</sup>/ ‘dancing girl’. In T4T1, the first tone shows a high-rising  $f_0$  (solid line in Figure 11(c)), as in /xəu<sup>4</sup> mɛ<sup>1</sup>/ ‘stepmother’. It is again very different from the high-falling  $f_0$  contour when the tone was pronounced in isolation (dashed line in Figure 11(c)).

Another aspect of tone sandhi worth noting is that in previous impressionistic studies, tone sandhi has always been argued to involve categorical changes from one lexical tone to another within the language’s tonal inventory, i.e. T1 + T1 → T3 + T1, T3 + T3 → T2 + T3, and T4 + T1 → T2 + T1 (e.g. Li & Liu 1985, Hung 1987, Tan 1987, Zhang 1987, Chen 2000, Wang 2002, Hyman 2007). This view, however, needs to be rectified.



**Figure 12** f0 realization of three tone sandhi patterns in Tianjin Mandarin compared to their purported sandhi output tone patterns. Solid lines for the f0 realization of the sandhi patterns, dashed lines for the purported sandhi output tone patterns. Time is normalized.

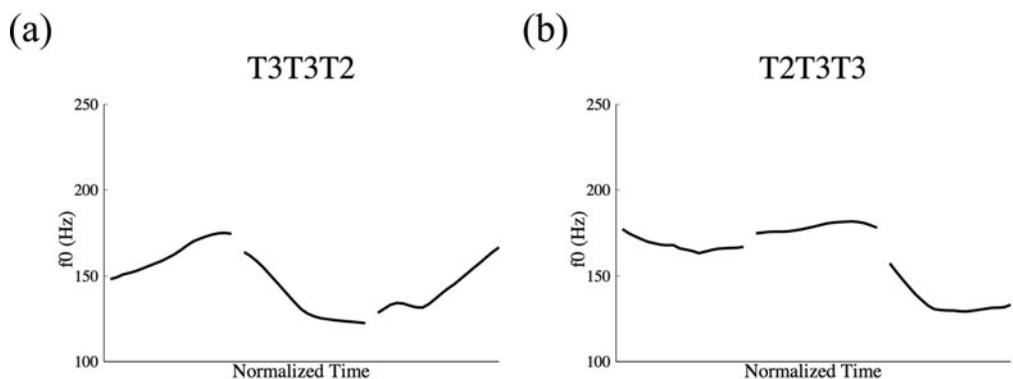
As shown in Figure 12, among the three tone sandhi sequences, T3T3 (solid line in Figure 12(b)) is near-merged with the claimed sandhi output tone sequence T2T3 (dashed line in Figure 12(b); as in /t<sup>h</sup>wo<sup>2</sup> nja<sup>3</sup>/ ‘ostrich’). T1T1 (solid line in Figure 12(a)) and T4T1 (solid line in Figure 12(c)), however, are far from their purported sandhi-derived lexical tonal contours, i.e. T3T1 (dashed line in Figure 12(a); as in /næ<sup>3</sup> mɛ<sup>1</sup>/ ‘nanny’) and T2T1 (dashed line in Figure 12(b); as in /pæ<sup>2</sup> mɑ<sup>1</sup>/ ‘white cat’), respectively. This argues strongly against the view that sandhi variations involve the change of one lexical tone to another. For more experimental data on disyllabic tone sandhi in Tianjin Mandarin, see Zhang & Liu (2011) and Li & Chen (2016).

Note that the disyllabic tone sandhi patterns plotted in Figures 11 and 12 (as well as in Li & Chen 2016) are based on speech of the young generation. Tianjin tone sandhi has been reported to undergo several diachronic changes. For example, Shi & Wang (2004) noted that T4T4 sandhi is applied among older speakers although it is no longer observed among middle-aged and young speakers. Liu & Gao (2003) also reported T4T4 sandhi as ‘obsolete’. In contrast, T4T1 sandhi seems to be a new innovation and could only be observed among middle-aged and young speakers; no T4T1 sandhi has been observed among old speakers (Shi & Wang 2004). Lu (1997) and Shi & Wang (2004) reported a high-level f0 realization of the first T1 of T1T1 among young speakers, which, however, is not observed in our dataset.

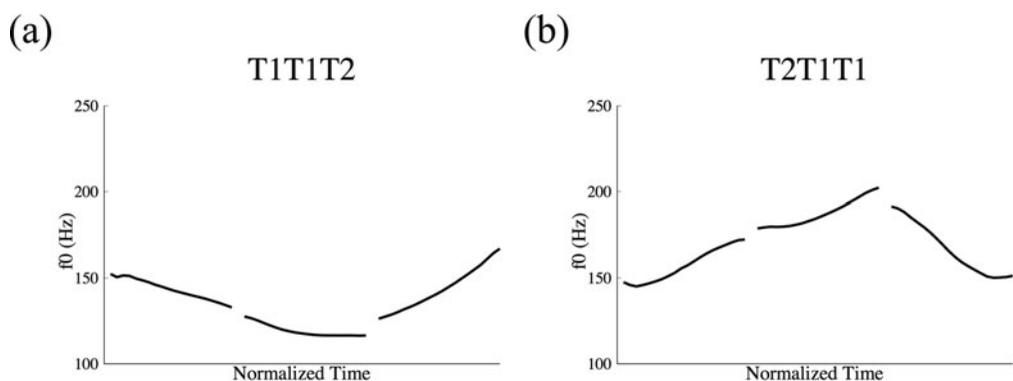
When these disyllabic tonal sequences occur in a larger domain such as trisyllabic sequences, tone sandhi has been claimed to apply consistently regardless of the alignment of the disyllabic sequences within a trisyllabic constituents (e.g. Li & Liu 1985, Chen 2000, Ma 2005, Wee et al. 2005). The purported consistent applications of disyllabic sandhi in trisyllabic sequences have given rise to much complexity in the analysis of sandhi applications, posing great challenges to theories of tonal alternation (e.g. Chen 2000, Yip 2002, Hyman 2007). For example, the pattern T1T1 has been claimed to apply when it is both left-aligned (e.g. T1T1T2) and right-aligned (e.g. T2T1T1). This, again, is not supported by our data based on speech of the younger generation.

Among the three tone sandhi sequences, T3T3 is realized with its sandhi change consistently and regardless of its alignment in the trisyllabic sequences, as illustrated in Figure 13. In both T3T3T2 (Figure 13(a); as in /xwo<sup>3</sup> pɛ<sup>3</sup> tɕje<sup>2</sup>/ ‘torch festival’) and T2T3T3 (Figure 13(b); as in /xwɛ<sup>2</sup> pja<sup>3</sup> tɕjaŋ<sup>3</sup>/ ‘Marble Pillar Award’), the first T3 is realized with a rising f0 contour (comparable to that in Figure 11(b)), indicating the application of tone sandhi in both contexts.

By contrast, T1T1 sandhi and T4T1 sandhi are only applied when the patterns are right-aligned in the trisyllabic sequences, as shown in Figures 14 and 15, respectively. In T2T1T1 (Figure 14(b)) where the T1T1 sequence is right-aligned as in /xuŋ<sup>2</sup> tɕŋ<sup>1</sup> tɕ<sup>h</sup>y<sup>1</sup>/ ‘red light district’, the middle T1 is realized with a rising f0 contour comparable to that in Figure 11(a), suggesting the application of tone sandhi in this case. When T1T1 is left-aligned (Figure 14(a)) as in /kuŋ<sup>1</sup> ʃaŋ<sup>1</sup> tɕy<sup>2</sup>/ ‘Trade and Industry Bureau’, tone sandhi does not apply since given



**Figure 13** f0 realization of trisyllabic sequences when T3T3 is left-aligned ((a); T3T3T2) vs. right-aligned ((b); T2T3T3). Time is normalized.



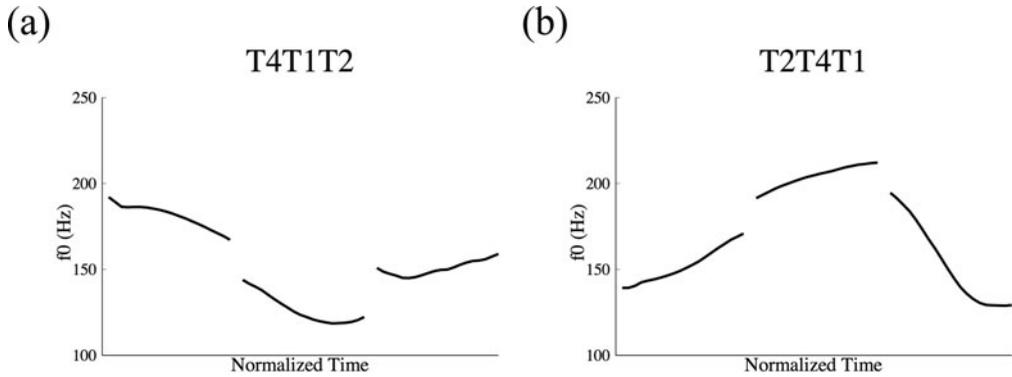
**Figure 14** f0 realization of trisyllabic sequences when T1T1 is left-aligned ((a); T1T1T2) vs. right-aligned ((b); T2T1T1). Time is normalized.

that the first T1 is realized with a falling f0 contour, comparable to its canonical form as in Figure 9. Similarly, right-aligned sandhi application could also be observed for the T4T1 sequence as in T2T4T1 (Figure 15(b); as in /tʂə<sup>2</sup> tɕi<sup>4</sup> pan<sup>1</sup>/ ‘acrobatic team’), while tone sandhi does not apply when the sequence is left-aligned as in T4T1T2 (Figure 15(a); as in /tʰæ<sup>4</sup> k<sup>h</sup>uŋ<sup>1</sup> tʃ<sup>h</sup>wan<sup>2</sup>/ ‘spaceship’). For more experimental data on trisyllabic tone sandhi in Tianjin Mandarin, see Li & Chen (2016).

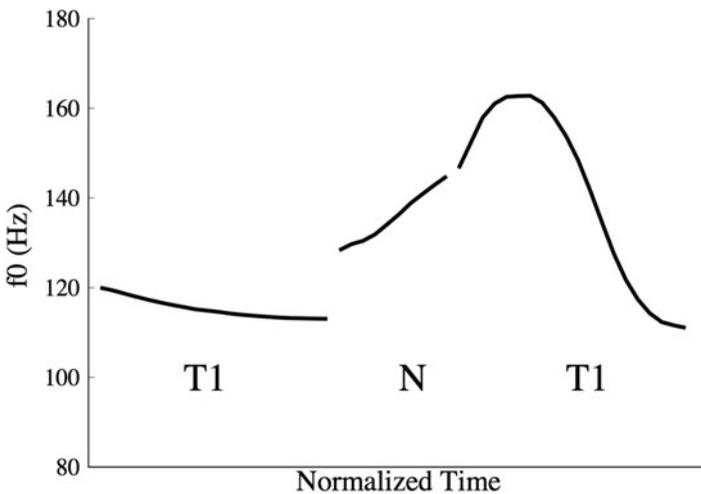
### Neutral tone

As in many other varieties of Mandarin Chinese, neutral tone also exists in Tianjin Mandarin. The neutral-tone syllables in Mandarin Chinese typically do not surface with any of the lexical tones (Chen 2015). As these syllables always occur in the prosodically weak positions (Chen & Xu 2006), they are usually produced with acoustic reduction at the segmental level, where the onset consonant of the neutral-tone syllable is sometimes voiced, and the vowel might be centralized or even deleted. For example, in the word /kʰ<sup>1</sup> kʰ/ ‘elder brother’, in which the second syllable is a neutral-tone syllable, its onset consonant is often voiced, and the vowel can be reduced to a schwa, realized as [kʰ<sup>1</sup> gə]. Neutral-tone syllables are usually produced with short duration (typically about half of the duration for a full lexical tone syllable) and their f0 realization also exhibits much variability.

Neutral-tone syllables in Tianjin Mandarin are common in grammatical morphemes (e.g. possessive marker /tə/ in /wo<sup>3</sup> tə/ ‘mine’), lexical items (e.g. /li/ in /pwo<sup>1</sup> li/ ‘glass’),



**Figure 15** f0 realization of trisyllabic sequences when T4T1 is left-aligned ((a); T4T1T2) vs. right-aligned ((b); T2T4T1). Time is normalized.

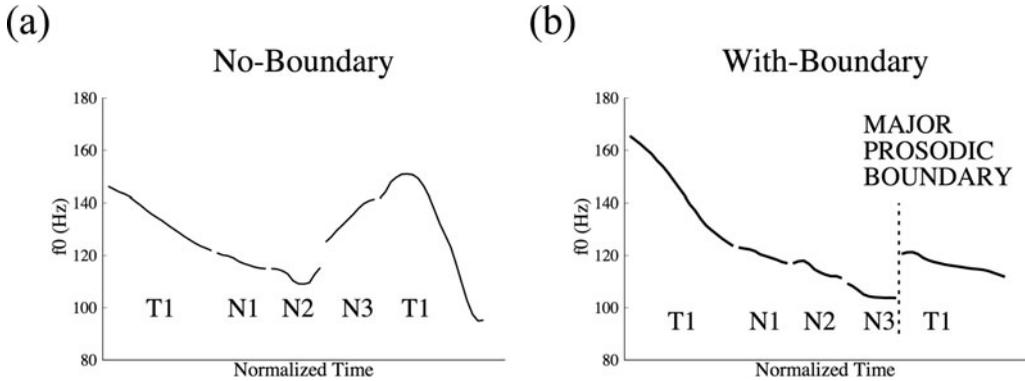


**Figure 16** f0 realization of one neutral tone (N) embedded between two T1s. Time is normalized.

diminutive words (e.g. /kʰ/ in /kʰʌ¹ kʰ/ ‘elder brother’), and reduplication (e.g. /kʰaŋ/ in /kʰaŋ⁴ kʰaŋ/ ‘to take a look’).

The f0 realization of neutral tone in Tianjin Mandarin is influenced by the preceding tone and shows varied patterns (e.g. Wang 2002, Li & Chen 2011). In particular, neutral tone before the low-falling T1 is very often realized with a rising f0 contour, typically when there is only one neutral tone embedded between two full tones (as in /peɪ¹ tʃə mə¹/ ‘carrying Mom on the back’). This is clearly shown with the rising f0 contour over the middle neutral-tone syllable (N) in Figure 16, in which each tonal contour was obtained by averaging across six samples produced by our speaker. This has led to the proposal that neutral tone in Tianjin Mandarin has a special high offset tonal target before the lexical T1, different from the typical low neutral-tone offset target before other lexical tones in the language (Wang 2002).

However, experimental data suggest that the rising neutral-tone realization could be due to the general raising effect of T1 upon its preceding tones (Li & Chen 2016). When there are multiple neutral tones like those shown in Figure 17(a), the mid-low neutral-tone target is approached first by the end of the second neutral tone (N2 in Figure 17(a)), as in the example sentence /tʰe¹ wo¹ mə¹ mə məŋ tə məɔ¹ tʰi¹ lwan⁴ lə nɛ⁴ kə ɕjeŋ⁴ tɕʰjəu²/ ‘He said mothers’ cats messed up that cotton ball’. The raised f0 realization of neutral tone could



**Figure 17** f0 realization of three neutral tones (N1N2N3) embedded between two T1s without (a) or with (b) a major prosodic boundary following the neutral tone sequence. Time is normalized.

only be observed over the very last neutral-tone syllable (N3 in Figure 17(a)). Importantly, the raising effect can be blocked by a major prosodic boundary as in Figure 17(b), as in the example sentence /tʰv¹ fwo¹ mɛ¹ mɛ mən̩ tɔ tɕjɛ¹ lə ʂan¹ pæ³ kʰwæ⁴ tɕʰjɛn̩²/ ‘He said mothers’ had increased by 300 yuan’.

**Transcription of recorded passage ‘North Wind and the Sun’**

The passage is transcribed phonemically with symbols described in the consonant and vowel sections. Full lexical tones are marked with superscript tone numbers instead of tonal values. As we do not follow the tradition that sandhi tones are transcribed as some other tone within the inventory, sandhi tones are marked with parentheses outside the original superscript tone numbers, e.g. <sup>(1)</sup> for sandhi-T1. Neutral-tone syllables are not marked with tone numbers. Syllable boundaries are marked by space, | marks the end of major phrases, and || for the end of utterances.

jəʊ³ ji<sup>(4)</sup> tʰjɛn̩¹ | pei³ fən̩¹ xɤ² tʰaɛ⁴ jan | tʃən̩⁴ tɕaɛ⁴ fwo¹ | ʃei² tɔ | pən̩³ ʃɿ tɕ³  
 || tʃən̩⁴ xɑɔ³ | jəʊ³ kə | tʃʰwan̩¹ tɔ³ pʰən̩ tɔ ɹən̩² | tɕəʊ³ kwo⁴ laɛ || tʰv¹ ljɛ³  
 fwo¹ | ʃei² n̩ən̩² ɹan̩⁴ | nei⁴ kə ɹən̩² | tʰwo¹ tjaɔ | tɔ³ pʰən̩ | tɕjəʊ⁴ ʂwan̩⁴ | ʃei²  
 li⁴ xae | ɹan̩² xəʊ⁴ | pei³ fən̩¹ tɕjəʊ⁴ | pʰin̩¹ lə min̩⁴ tɔ tʃʰwei¹ || tan̩⁴ ʃɿ | tʰv¹  
 tʃʰwei¹ tɔ | ɤ⁴ li⁴ xae | nei⁴ kə ɹən̩² | tɕjəʊ⁴ pɛ<sup>(3)</sup> tɔ³ pʰən̩ | kwo³ tɔ ɤ⁴ tɕin̩³  
 || pei³ fən̩¹ | mei² tʃɤ² | tʃɿ<sup>(3)</sup> xɑɔ³ | fan̩⁴ tɕhi⁴ || tɕje¹ tʃə | tʰaɛ⁴ jan | ʃaɛ⁴ tɔ |  
 ɹɤ⁴ xun̩<sup>(1)</sup> xun̩¹ tɔ | nei⁴ kə ɹən̩² | li⁴ mɛ<sup>(3)</sup> pɛ³ | tɔ³ pʰən̩ | tʰwo¹ tjaɔ⁴ lə || ɹan̩²  
 xəʊ⁴ | pei³ fən̩¹ | tʃɿ³ n̩ən̩² | ɹən̩<sup>(4)</sup> ʃu¹ ||

**Orthographic Transcription**

有一天，北风和太阳正在说谁的本事大。正好有个穿斗篷的人走过来。他俩说，谁能让那个人脱掉斗篷，就算谁厉害。然后，北风就拼了命地吹。但是他吹得越厉害，那个人就把斗篷裹得越紧。北风没辙，只好放弃。接着，太阳晒得热烘烘的，那个人立马把斗篷脱掉了。然后，北风只能认输。

**Acknowledgements**

This research was funded by China Postdoctoral Science Foundation and China Scholarship Council (Qian Li), as well as the ERC Starting grant and the KNAW–China Exchange grant (Yiya Chen).

We thank the editors Amalia Arvaniti and Adrian Simpson, our copyeditor Ewa Jaworska, and two anonymous reviewers for their constructive suggestions. We also thank Shipeng Shao for stimuli recording. Thanks to Wen Cao and Feng Shi for making the recording booths available. We are grateful to Menghui Shi for sharing scripts. We also benefit greatly from the inspiring discussions with Ting Zeng.

## References

- Boersma, Paul & David Weenink. 2017. Praat: Doing phonetics by computer (version 6.0.16). <http://www.fon.hum.uva.nl/praat/> (accessed 18 Jan 2017).
- Chao, Yuen-Ren. 1920. A system of tone-letters. *Le Maître Phonétique* 45, 24–27.
- Chao, Yuen-Ren. 1948. *Mandarin primer*. Cambridge, MA: Harvard University Press.
- Chen, Matthew Y. 2000. *Tone sandhi: Patterns across Chinese dialects*. Cambridge: Cambridge University Press.
- Chen, Yiya. 2015. Neutral tone. In Rint Sybesma (ed.), *Encyclopedia of Chinese language and linguistics*, [http://dx.doi.org/10.1163/2210-7363\\_cell\\_COM\\_000162](http://dx.doi.org/10.1163/2210-7363_cell_COM_000162) (accessed 18 January 2017). Leiden: Brill.
- Chen, Yiya & Carlos Gussenhoven. 2015. Shanghai Chinese. *Journal of the International Phonetic Association* 45, 321–337.
- Chen, Yiya & Yi Xu. 2006. Production of weak elements in speech: Evidence from f0 patterns of neutral tone in Standard Chinese. *Phonetica* 63, 47–75.
- Han, Gendong. 1993a. Tianjin fangyan de yuyin fenxi [A phonetic analysis of Tianjin Mandarin]. In Han (ed.), 15–26.
- Han, Gendong. 1993b. Tianjinren zenyang xue putonghua [How do Tianjiners learn Standard Chinese]. In Han (ed.), 284–304.
- Han, Gendong (ed.). 1993c. *Tianjin fangyan* [Tianjin Mandarin]. Beijing: Yanshan Chubanshe.
- Huang, C.-T. James, Y.-H. Audrey Li & Andrew Simpson (eds.). 2014. *The handbook of Chinese linguistics*. Chichester: Wiley Blackwell.
- Hung, Tong-Ning. 1987. Tianjin tone sandhi: Towards a unified approach. *Journal of Chinese Linguistics* 15, 274–305.
- Hyman, Larry M. 2007. Universals of tone rules: 30 years later. In Tomas Riad & Carlos Gussenhoven (eds.), *Tones and tunes 1: Typology studies in word and sentence prosody*, 1–34. Berlin: Mouton de Gruyter.
- Karlgren, Bernhard. 1915–1926. *Études sur la phonologie Chinoise* [Archives d'Études Orientales]. Leiden & Stockholm: E. J. Brill.
- Ladefoged, Peter. 1999. American English. In IPA (ed.), *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet*, 41–44. Cambridge: Cambridge University Press.
- Ladefoged, Peter. 2006. *A course in phonetics*. London: Wadsworth Publishing.
- Ladefoged, Peter & Ian Maddieson. 1996. *The sounds of the world's languages*. Oxford: Blackwell.
- Ladefoged, Peter & Zongji Wu. 1984. Places of articulation: An investigation of Pekingese fricatives and affricates. *Journal of Phonetics* 12, 267–278.
- Lee, Wai-Sum & Eric Zee. 2003. Standard Chinese (Beijing). *Journal of the International Phonetic Association* 33, 109–112.
- Lee, Wai-Sum & Eric Zee. 2014. Chinese phonetics. In Huang et al. (eds.), 369–399.
- Lee-Kim, Sang-Im. 2014. Revisiting Mandarin 'apical vowels': An articulatory and acoustic study. *Journal of the International Phonetic Association* 44, 261–282.
- Li, Qian & Yiya Chen. 2011. Neutral tone realization in Tianjin Mandarin. In Wai-Sum Lee & Eric Zee (eds.), *17th International Congress of Phonetic Sciences (ICPhS XVII)*, Hong Kong, 1214–1217.
- Li, Qian & Yiya Chen. 2016. An acoustic study of contextual tonal variation in Tianjin Mandarin. *Journal of Phonetics* 54, 123–150.
- Li, Xingjian & Sixun Liu. 1985. Tianjin fangyan de liandu biandiao [Tone sandhi in Tianjin Mandarin]. *Zhongguo Yuwen* 1, 76–80.
- Lin, Yen-Hwei. 2007. *Sounds of Chinese*. Cambridge: Cambridge University Press.

- Lin, Yen-Hwei. 2014. Segmental phonology. In Huang et al. (eds.), 400–421.
- Liu, Sixun. 1993. Tianjin fangyan de yuyin biandiao [Tone sandhi in Tianjin Mandarin]. In Han (ed.), 284–294.
- Liu, Yuzhen & Jing Gao. 2003. Qu–qu liandu biandiao guize: Shehui yuyanxue bianxiang [T4–T4 tone sandhi rule: Sociolinguistic variation]. *Tianjin Shifan Daxue Xuebao (Shehui Kexue Ban)* 5, 65–69.
- Lu, Jilun. 1997. Tianjin fangyan zhong de yizhong xinde liandu biandiao [A new tone sandhi rule in Tianjin Mandarin]. *Tianjin Shifan Daxue Xuebao* 4, 67–72.
- Lu, Jilun (ed.). 2004. *Phonetic and phonological studies on Tianjin dialect*. Beijing: Beijing Ligong Daxue Press.
- Ma, Qiuwu. 2005. ‘Tianjinhua liandu biandiao zhi mi’ de youxuanlun jieshi [An optimality account of ‘the paradox of Tianjin tone sandhi’]. *Zhongguo Yuwen* 6, 561–568.
- Ma, Qiuwu & Yuan Jia. 2006. Tianjinhua shangsheng de liangtiao ‘biandiao guize’ bianxi [An analysis of two ‘sandhi rules’ of T3 in Tianjin Mandarin]. *Tianjin Shifan Daxue Xuebao (Shehui Kexue Ban)* 1, 53–58.
- Shi, Feng. 1986. Tianjin fangyan shuangzizu shengdiao fenxi [An analysis of disyllabic tonal combination in Tianjin Mandarin]. *Yuyan Yanjiu* 1, 77–90.
- Shi, Feng & Ping Wang. 2004. Tianjinhua shengdiao de xinbianhua [New change of tones in Tianjin Mandarin]. In Feng Shi & Zhongwei Shen (eds.), *Lezaiqizhong – Wangshiyuan jiaoshou qishi huadan qingzhu wenji* [The joy of research: A Festschrift in honor of Professor William S.-Y. Wang on his seventieth birthday], 176–188. Tianjin: Nankai University Press.
- Tan, Fu. 1987. Tone sandhi in the Tianjin dialect. *Journal of Chinese Linguistics* 15, 228–246.
- Wang, Jialing. 2002. Youxuanlun he tianjinhua de liandu biandiao ji qingsheng [Optimality Theory and tone sandhi and neutral tone in Tianjin Mandarin]. *Zhongguo Yuwen* 4, 363–372.
- Wang, Jialing & Hui Jiang. 1997. Tianjinhua qingsheng de yuyin xingzhi he yinxi fenxi [The phonetic properties and phonological analysis of neutral tone in Tianjin Mandarin]. *Qinzhu zhongguo shehui kexueyuan yuyan yanjiusuo jiansuo sishiwu zhounian xueshu lunwenji* [Collected works in honor of the 45th anniversary of Institute of Linguistics at Chinese Academy of Social Sciences]. Beijing: Shangwu Press.
- Wang, Zhijie. 1997. Yinghan yinjie biyunwei de butong xingzhi [Different natures of nasal codas in English and Chinese]. *Xiandai Waiyu* 4, 17–29.
- Wee, Lian-Hee, Xiuhong Yan & Matthew Y Chen. 2005. *Yinan yu Luxiang – Lun Tianjin Fangyan de Liandu Biandiao* [Problems and solutions: On tone sandhi in Tianjin Mandarin]. Beijing: Shangwu Press.
- Yip, Moira. 2002. *Tone*. Cambridge: Cambridge University Press.
- Zhang, Jie & Jiang Liu. 2011. Tone sandhi and tonal coarticulation in Tianjin Chinese. *Phonetica* 68, 161–191.
- Zhang, Zheng-Sheng. 1987. The paradox of Tianjin: Another look. *Journal of Chinese Linguistics* 15, 228–246.