Tianjin Mandarin

Qian Li[a,b]* q.li@hum.leidenuniv.nl
Yiya Chen[b,c] yiya.chen@hum.leidenuniv.nl
Ziyu Xiong[a] xiongziyu@163.com

[a] Institute of Linguistics, Chinese Academy of Social Sciences, 100732, Beijing, China
[b] Leiden University Center for Linguistics, Postbus 9515, 2300 RA, Leiden, the Netherlands
[c] Leiden Institute for Brain and Cognition, Postbus 9515, 2300 RA, Leiden, the Netherlands
* Corresponding author. Tel.: +86 10 8519 5394

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, 1993b; 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 has 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., [1] for

Tone 1, ˧˥ for Tone 2, ˨˩˦ for Tone 3 and ˥˩ for Tone 4. (See Section 'Lexical Tones' for more details on lexical tones in Tianjin Mandarin.)

Consonants

	Bilabial	Labiodental	Dentoalveolar	Postalveolar	Palatal	Alveolo-palatal	Velar
Plosive	p pʰ		t̪ t̪ʰ				k kʰ
Affricate			ts̪ ts̪ʰ	tʃ̪ tʃ̪ʰ		tɕ tɕʰ	
Nasal	m		n̪				ŋ
Fricative		f	s̪	ʃ̪		ɕ	x
Approximant	w			ɻ̪	j ɥ		
Lateral			l̪				

	PHONETIC	ORTHOGRAPHIC	GLOSS
p	pɐ˧	八	'eight'
pʰ	pʰɐ˨˩˦	爬	'to climb'
m	mɐ˧	妈	'mother'
w	wɐ˧	蛙	'frog'
f	fɐ˧	发	'to send'
t̪	t̪ɐ˧	搭	'to build'
t̪ʰ	t̪ʰɐ˧	他/她	'he/she'
ts̪	ts̪ɐ˧	匝	'to circle'
ts̪ʰ	ts̪ʰɐ˧	擦	'to wipe'
n̪	n̪ɐ˥˩	纳	'to include'
s̪	s̪ɐ˧	撒	'to cast'
l̪	l̪ɐ˧	拉	'to pull'
tʃ̪	tʃ̪ɐ˧	渣	'residue'

tʃ̪ʰ	tʃ̪ʰɐ¹	插	‘to insert’
ʃ̪	ʃ̪ɐ¹	沙	‘sand’
ɻ̪	ɻ̪ən̪²	人	‘person’
tɕ	tɕjɐ¹	加	‘to add’
tɕʰ	tɕʰjɐ¹	掐	‘to nip off’
ɕ	ɕjɐ¹	虾	‘shrimp’
j	jɐ¹	鸭	‘duck’
ɥ	ɥe¹	约	‘to restrict’
k	kɤ¹	歌	‘song’
kʰ	kʰɤ¹	科	‘subject’
ŋ	ɑŋ²	昂	‘to raise’
x	xɤ¹	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.)

Plosive

Plosives in Tianjin Mandarin differentiate three places of articulation: bilabial /p, pʰ/ as in /pɐ¹/ ‘eight’ and /pʰɐ²/ ‘to climb’, dentoalveolar /t̪, t̪ʰ/ as in /t̪ɐ¹/ ‘to build’ and /t̪ʰɐ¹/ ‘he/she’, and velar /k, kʰ/ as in /kɤ¹/ ‘song’ and /kʰɤ¹/ ‘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.

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

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.

Affricate

Affricates in Tianjin Mandarin display the same two-way distinction in aspiration as plosives. They have three places of articulation: dentoalveolar /ʦ̻, ʦ̻^{h}/ as in /ʦ̻ɐ^{1}/ ‘to circle’ and /ʦ̻^{h}ɐ^{1}/ ‘to wipe’, postalveolar /ʧ̻, ʧ̻^{h}/ as in /ʧ̻ɐ^{1}/ ‘residue’ and /ʧ̻^{h}ɐ^{1}/ ‘to insert’, as well as alveolo-palatal /ʨ, ʨ^{h}/ as in /ʨjɐ^{1}/ ‘to add’ and /ʨ^{h}jɐ^{1}/ ‘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 there are some word-specific generational differences in the specific place of articulation for this series of sounds. As observed in Han (1993a, 1993b), while the young-variety Tianjin Mandarin speakers produce some words using the postalveolar consonants (i.e., /tʃ̻, tʃ̻^{h}, ʃ̻/), older generation speakers typically produce them with their dentoalveolar counterparts (i.e., /ts̻, ts̻^{h}, s̻/, respectively) (cf. 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 /tɕjɐ^{1}/ 'to add' and /tɕ^{h}jɐ^{1}/ 'to nip off'. Figure 1 shows the spectrogram of /tɕjɐ^{1}/ 'to add', where we see a transition (glide /j/) between the consonant /tɕ/ and the vowel /ɐ/, taking up about 1/4 ~ 1/3 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 /ɐ/; the F2 of the glide starts around 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 /tɕia^{1}/ and /tɕ^{h}ia^{1}/ (e.g., Wee et al., 2005). In line with Lin (2007), however, we treat /j/ as a glide which constitutes part of the onset.

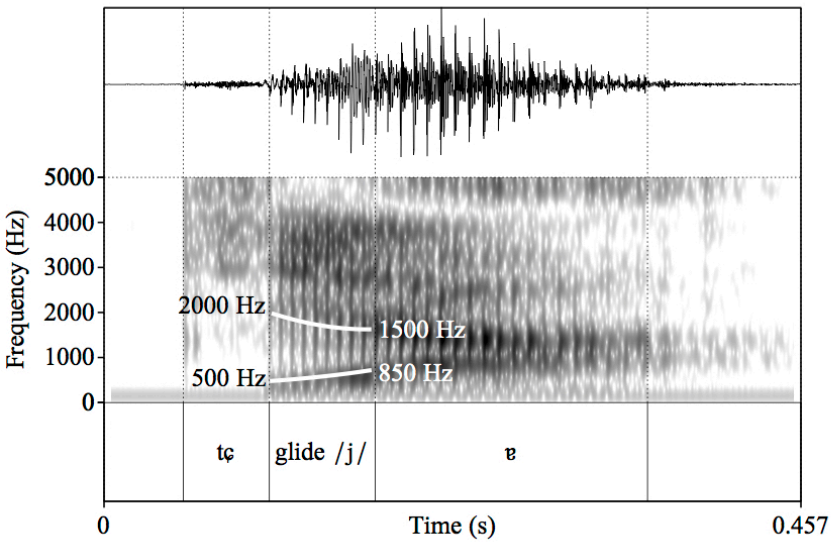


Figure 1 *The spectrogram of the whole syllable /tɕjɐ^{1}/.*

Figure 2 plots the spectrograms of three monomorphemic words with the same rhyme: /t̪ɑɔ^{1}/ ‘knife’, /t̪jɑɔ^{1}/ ‘to hold in the mouth’, and /tɕjɑɔ^{1}/ ‘to teach’. There is a clear glide-like transition in /tɕjɑɔ^{1}/ (Figure 2(c)), which is similar to that in /t̪jɑɔ^{1}/ (Figure 2(b)) where there is a real glide. Both are different from that in /t̪ɑɔ^{1}/ (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.

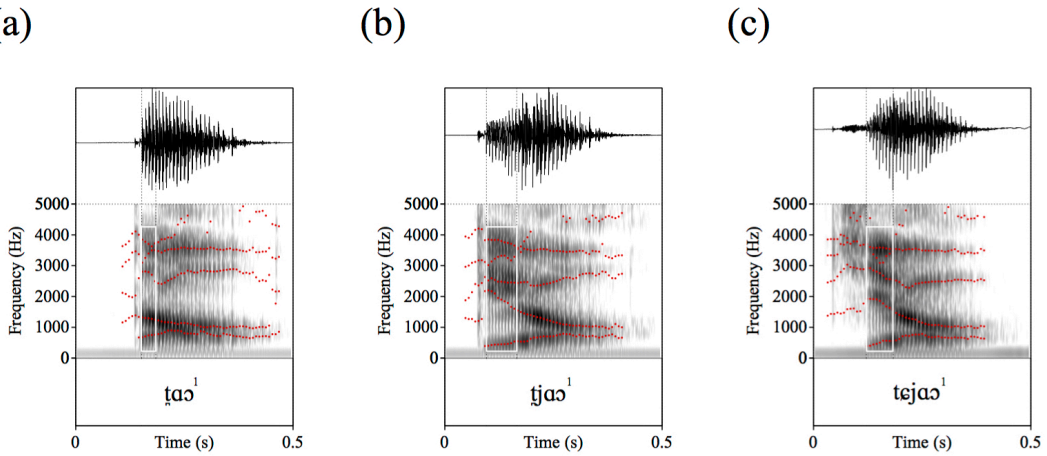


Figure 2 *The spectrograms of /t̪ɑɔ^{1}/ (a), /t̪jɑɔ^{1}/ (b), and /tɕjɑɔ^{1}/ (c). White frames indicate the transitions.*

Nasal

Tianjin Mandarin has nasals in three places of articulation: bilabial /m/ as in /mɤ^{1}/ ‘mother’, dentoalveolar /n̪/ as in /n̪ɤ^{4}/ ‘to include’, and velar /ŋ/ as in /ɑŋ^{2}/ ‘to raise’. /m/ can only occur in the onset position while /ŋ/ only in the coda position.

Fricative

Fricatives in Tianjin Mandarin differentiate five places of articulation: labiodental /f/ as in /fɐ^{1}/ ‘to send’, dentoalveolar /s̪/ as in /s̪ɐ^{1}/ ‘to cast’, postalveolar /ʃ̪/ as in /ʃ̪ɐ^{1}/ ‘sand’, alveolo-palatal /ɕ/ as in /ɕjɐ^{1}/ ‘shrimp’, and velar /x/ as in /xɤ^{1}/ ‘to drink’. The alveolo-palatal fricative in Tianjin Mandarin is obligatorily followed by a palatal glide (as in /ɕjɐ^{1}/ ‘shrimp’ and /ɕɥe^{1}/ ‘boots’) or a high vowel (as in /ɕi^{1}/ ‘west’). (See Section ‘Lateral and Approximant’ 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 /xaɛ^{2}/ ([χaɛ^{2}]) ‘child’ and /xɑɔ^{3}/ ([χɑɔ^{3}]) ‘good’.

Lateral and Approximant

Tianjin Mandarin has one lateral /l̪/ as in /l̪ɐ^{1}/ ‘to pull’ and four approximants: /w/, /ɻ̪/, /j/, and /ɥ/, as in /wɐ^{1}/ ‘frog’, /ɻ̪ən̪^{2}/ ‘person’, /jɐ^{1}/ ‘duck’, and /ɥe^{1}/ ‘to restrict’. Among the four approximants, /w/, /j/ and /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, which mainly contrast in lip rounding. They do not occur in the same contexts except before the vowel /e/, as in /n̪je^{1}/ ‘to pinch’ vs. /n̪ɥe^{4}/ ‘to abuse’, /tɕje^{1}/ ‘to connect’ vs. /tɕɥe^{2}/ ‘to feel’, /tɕ^{h}je^{1}/ ‘to cut’ vs. /tɕ^{h}ɥe^{1}/ ‘to lack’, and /ɕje^{1}/ ‘to rest’ vs. /ɕɥe^{1}/ ‘boots’. According to Han (1993a, 1993b), as also pointed out by one of the reviewers, the onset /ɻ̪/ can be pronounced as [j] in the old variety of Tianjin Mandarin in words such as /ɻ̪əu^{4}/ ([jəu^{4}]) ‘meat’; and interestingly, the onset /j/ can also be pronounced as [ɻ̪] in words such as /jɑŋ^{4}/ ([ɻ̪ɑŋ^{4}]) ‘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 Consonant

Tianjin Mandarin has two syllabic consonants (notation in line with Chao, 1948): the dentoalveolar /ɻ̩/ (as in /ts̪ɻ̩^{3}/ ‘son’) and postalveolar /ʐ̩/ (as in /tʃ̪ʐ̩^{3}/ ‘paper’). These two sounds

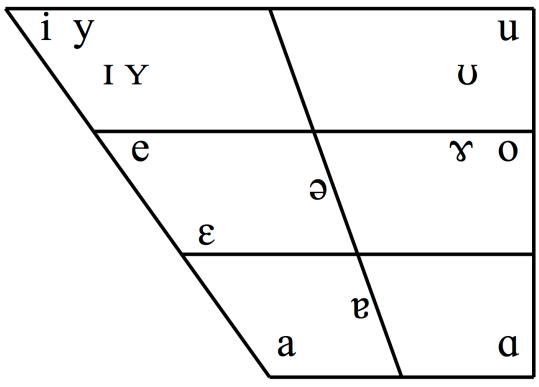
are traditionally referred to as ‘apical vowels’ by Sinologists (e.g., Karlgren, 1915–1926). Ladefoged & Maddieson (1996) refer them to as ‘fricative vowels’. With evidence from both ultrasound imaging and acoustic data, Lee-Kim (2014) shows that these two sounds are neither vowels nor fricative, but more comparable with approximants in nature, i.e., syllabic dental approximant /ɹ̪̩/ and retroflex approximant /ɻ̩/, respectively. This is similar with observations made in Lee & Zee (2003, 2014), in which, however, the two sounds have been transcribed with the same syllabic approximant /ɹ̩/.

The dentoalveolar /ɹ̪̩/ only follows dentoalveolar consonants /ts̪, tsʰ̪, s̪/ (as in /ts̪ɹ̪̩³/ ‘son’, /tsʰ̪ɹ̪̩²/ ‘word’, /s̪ɹ̪̩¹/ ‘to think’), while /ʐ̪̩/ follows postalveolar consonants /tʃ̪, tʃʰ̪, ʃ̪/ (as in /tʃ̪ʐ̪̩³/ ‘paper’, /tʃʰ̪ʐ̪̩¹/ ‘to eat’, /ʃ̪ʐ̪̩¹/ ‘poem’). Their tongue configurations are similar with the preceding homorganic consonants, i.e., /ɹ̪̩/ is homorganic with /ts̪, tsʰ̪, s̪/, and /ʐ̪̩/ is homorganic with /tʃ̪, tʃʰ̪, ʃ̪/. In addition, /ɹ̪̩/ can occur by itself as in /ɹ̪̩⁴/ ‘the sun’. To highlight the homoganicity 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 /ʐ̪̩/ symbol rather than the retroflex /ɻ̩/.

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

Monophthong



	PHONETIC	ORTHOGRAPHIC	GLOSS
i	t̪i^{1}	低	‘low’
y	l̪y^{2}	驴	‘donkey’
e	t̪je^{1}	爹	‘dad’
ɐ	t̪ɐ^{1}	搭	‘to build’
o	t̪wo^{1}	多	‘more’
ɤ	t̪ɤ^{2}	得	‘to get’
u	t̪u^{1}	都	‘metropolis’
ə	t̪ə	的	Possessive marker
	t̪ən̪^{4}	扽	‘to drag’
	t̪əŋ^{1}	灯	‘lamp’
ɪ	l̪ɪn̪^{2}	林	‘forest’
	l̪ɪŋ^{2}	零	‘zero’
ʏ	tɕʏn̪^{1}	军	‘army’
ɛ	t̪jɛn̪^{1}	颠	‘bump’
a	t̪an̪^{1}	单	‘single’
ɑ	t̪ɑŋ^{1}	当	‘when’
ʊ	t̪ʊŋ^{1}	东	‘east’

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

Among the seven vowels that occur only in open syllables, /i/ and /y/ (as in /t̪i^{1}/ ‘low’ and /l̪y^{2}/ ‘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 /t̪je^{1}/ ‘dad’. /ɐ/ is a low mid vowel as in /t̪ɐ^{1}/ ‘to build’. /u/ is a high back rounded vowel as in /t̪u^{1}/ ‘metropolis’. Both /ɤ/ and /o/

are mid-high back vowels differing mainly in lip rounding as in /t̪ɤ²/ ‘to get’ and /t̪wo¹/ ‘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 /l̪ɪn̪²/ ‘forest’ and /l̪ɪŋ²/ ‘zero’) and /ʏ/ (as in /tɕʏn̪¹/ ‘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 /t̪ʰɪŋ¹/ ([t̪ʰɪᵊŋ¹] ‘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 (as shown with an arrow in Figure 3(b)) without the presence of a transitional schwa. One of our reviewers pointed out, the nasal codas in Tianjin Mandarin are possibly nasal glides, while the English nasal codas are nasal stops (also see Wang, 1997), which we agree might be the reason behind the vowel–coda juncture difference. Further studies are needed to verify this possibility.

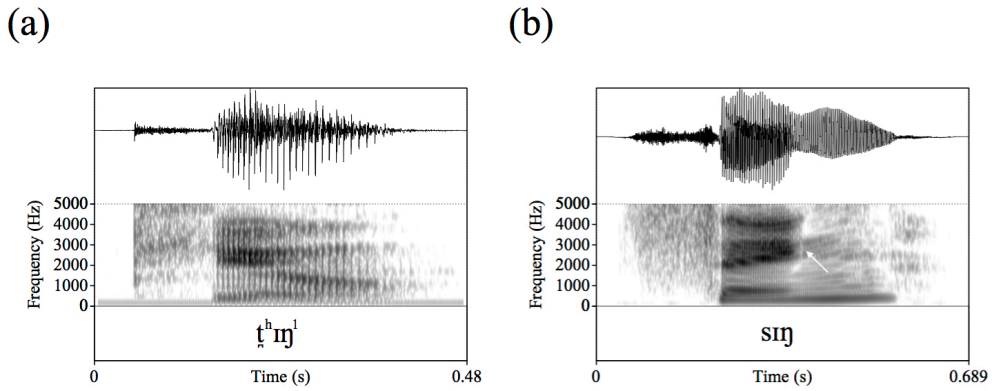


Figure 3 *Spectrogram comparison of the syllable /t̪ʰɪŋ¹/ in Tianjin Mandarin (a) vs. the syllable /sɪŋ/ in American English (b).*

/ɛ/ is a mid-low front vowel which only occurs before the dentoalvelar nasal coda as in /tjɛn̪¹/ ‘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 /t̪an̪[1]/ ‘single’), and the low back vowel /ɑ/ before the velar nasal coda (as in /t̪ɑŋ[1]/ ‘when’). /ʊ/ is the lax counterpart of /u/, which occurs before the velar nasal coda as in /t̪ʊŋ[1]/ ‘east’.

/ə/ is a central vowel which can be in both open and closed syllables although an open syllable with /ə/ is exclusively a neutral-tone syllable (e.g. /t̪ə/ *possessive marker* in /wo[3] t̪ə/ ‘mine’). (See section ‘Neutral Tone’ for more details on neutral tone.) In closed syllables, /ə/ can occur before both nasal codas /n̪/ and /ŋ/ (as in /t̪ən̪[4]/ ‘to drag’ and /t̪əŋ[1]/ ‘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 formula (following Boersma & Weenink, 2017):

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

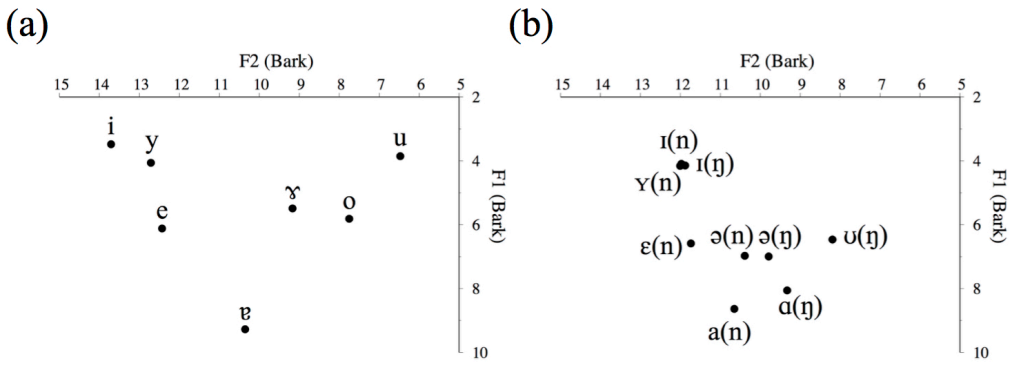
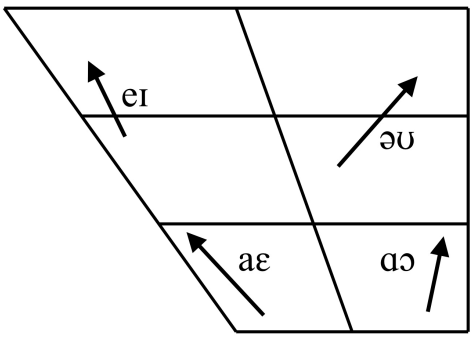


Figure 4 *F1 and F2 values (in Bark) of monophthongs produced in open syllables (a) and in closed syllables (b).*

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 /n̪/, 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^ha^2/ 'to climb') and closed syllables (e.g., /san[1]/ 'three' and /taŋ[3]/ 'party'). However, given the clearly different vowel quality, here we adopt an alternative view to treat them as different phonemes (also see Lin, 2007; 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).

Diphthong



	PHONETIC	ORTHOGRAPHIC	GLOSS
eɪ	l̪eɪ[2]	雷	'thunder'
aɛ	t̪aɛ[1]	呆	'dull'
ɑɔ	t̪ɑɔ[1]	刀	'knife'
əʊ	t̪əʊ[1]	都	'all'

There are four diphthongs in Tianjin Mandarin, with /eɪ, aɛ/ gliding towards the front (as in /l̪eɪ²/ 'thunder' and /t̪aɛ¹/ 'dull') and /ɑɔ, əʊ/ towards the back (as in /t̪ɑɔ¹/ 'knife' and /t̪əʊ¹/ '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 Hz to Bark using the formula (following Boersma & Weenink, 2017):

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

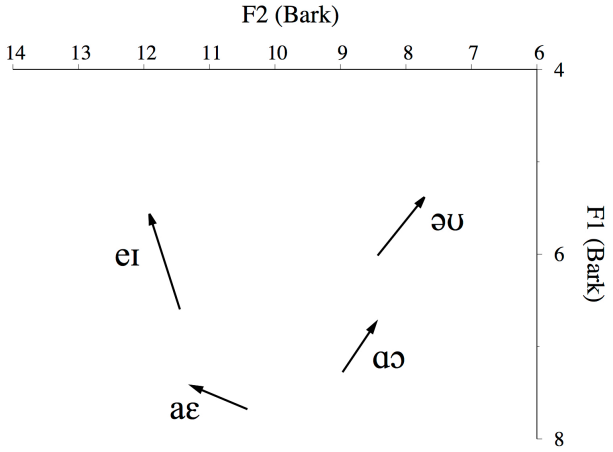


Figure 5 *F1 and F2 values (in Bark) of diphthongs. Gliding trajectories are shown with arrows.*

/eɪ/ and /aɛ/ are frequently transcribed as /ei/ and /ai/, respectively, both gliding towards the same high front target /i/ (e.g., Han, 1993a, 1993b; 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 ending points of /eɪ/ and /aɛ/, Figure 6 compares their spectrograms. As shown in Figure 6, the ending 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).

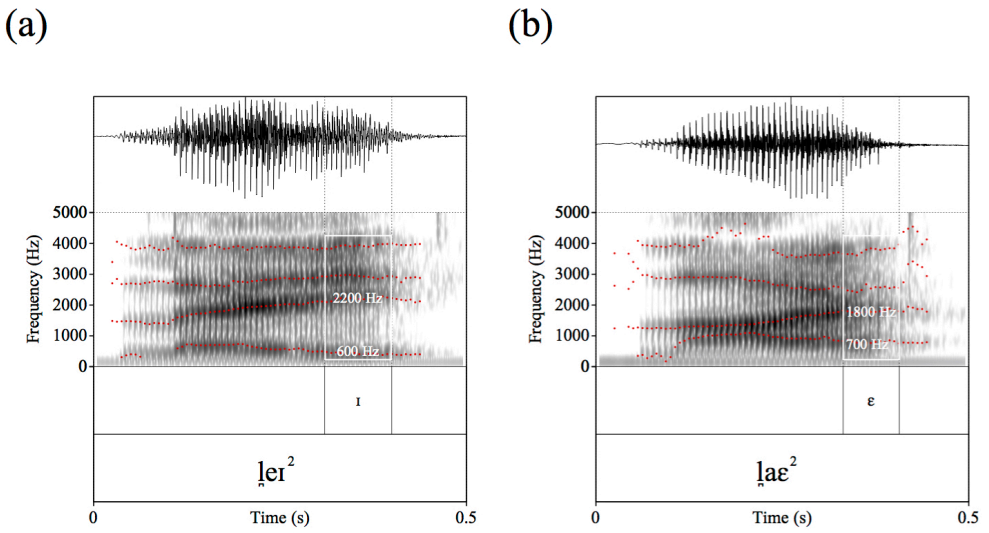


Figure 6 *Spectrogram comparison of diphthongs /eɪ/ vs. /aɛ/ as in syllables /l̪eɪ²/ 'thunder' (a) vs. /l̪aɛ²/ 'to come' (b), respectively.*

Similar differences can also be observed for /ɑɔ/ vs. /əʊ/, both of which have been frequently described to glide towards the high back vowel /u/ (e.g., Han, 1993a, 1993b; Wee et al., 2005). Figure 7 compares the spectrograms of /ɑɔ/ vs. /əʊ/, which again show different ending qualities of the two vowels. To be specific, /əʊ/ (Figure 7(b)) shows lower F1 (450 Hz) and F2 (850 Hz) than those of /ɑɔ/ (F1: 650 Hz; F2: 950 Hz; see Figure 7(a)).

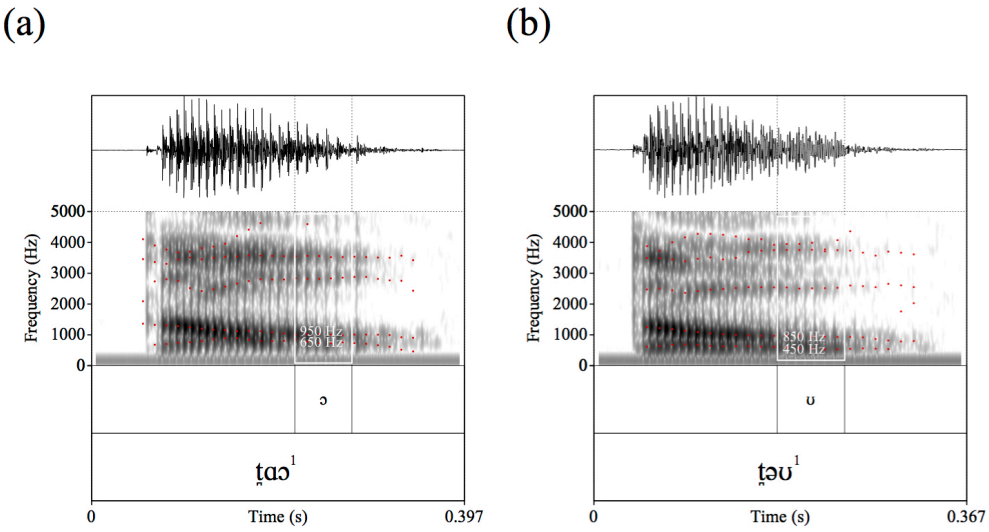


Figure 7 *Spectrogram comparison of diphthongs /ɑɔ/ vs. /əʊ/ as in syllables /t̪ɑɔ¹/ 'knife' (a) vs. /t̪əʊ¹/ 'all' (b), respectively.*

Tianjin Mandarin has a rhotic vowel /ɚ/ which is produced as an r-colored schwa, with the tip of tongue raised. /ɚ/ is syllabic as in /ɚ^{2}/ ‘son’, /ɚ^{3}/ ‘ear’ and /ɚ^{4}/ ‘two’. When /ɚ/ is produced in a neutral-tone syllable, it is used as a diminutive suffix.

When adding the suffix /ɚ/ to a noun, the two syllables are typically coalesced into one rhotacized syllable in the output form. The vowel part of the preceding syllable is directly rhotacized and only the lexical tone of the preceding syllable is kept, e.g., /t̪i^{3}/ + /ɚ/ → [t̪i˞^{3}] ‘remnants’ (as compared to [t̪i^{3}] ‘bottom’). Such process is known as ‘rhotacization’ or ‘er-hua’ in Chinese. Figure 8 compares the spectrograms of non-rhotacized [i] to its rhotacized counterpart [i˞]. As shown in Figure 8(b), the first 1/4 of [i˞] is realized similarly to [i] (as in Figure 8(a)). The remaining 3/4, however, is realized with a clearly lowered mean F3 (from about 2500 Hz to 1920 Hz), which is a typical acoustic cue of rhotacization (Ladefoged & Maddieson, 1996). In addition, the process of rhotacization 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, rhotacization has been traditionally transcribed as /ɚ/-insertion, e.g., /i/ + /ɚ/ → [iɚ] (see e.g., Han, 1993a). However, given the clear F3 lowering in the rhotacized vowel, we regard the F1 and F2 changes as the by-products of rhotacization rather than a rhotacized schwa insertion.

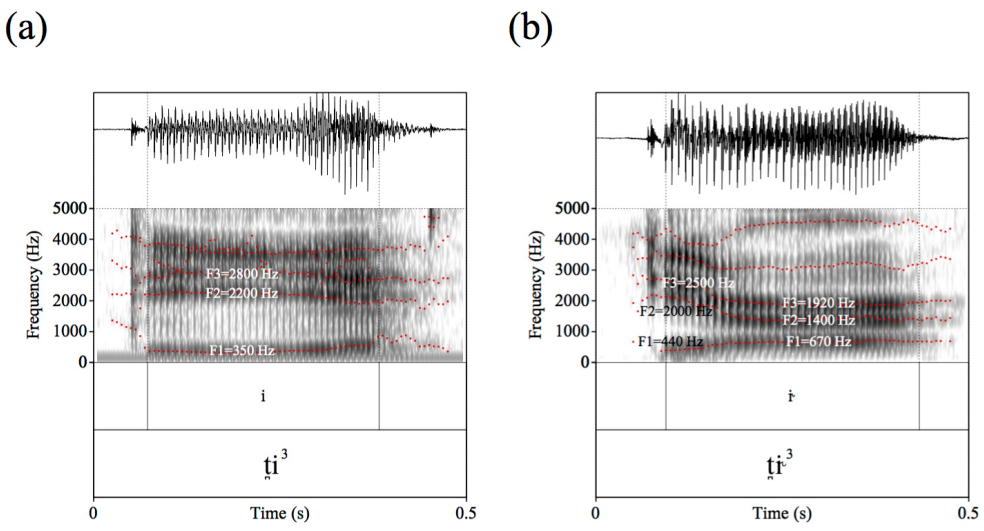


Figure 8 *Spectrogram comparison of non-rhotacized [i] vs. rhotacized [i˞] as in syllables [t̪i³] 'bottom' (a) vs. [t̪i˞³] 'remnants' (b), respectively.*

There are also other rhotacizing processes in both open and closed syllable structures. In open syllables, if the vowel is /aɛ/, only the /a/ part is rhotacized while /ɛ/ is deleted, e.g., /pʰaɛ²/ + /ɚ²/ → [pʰa˞²] 'badge' (as compared to [pʰaɛ²] 'card'); if the rhyme is /e/, /eɪ/ or a syllabic consonant, the entire rhyme part is replaced with /ɚ/, e.g., /peɪ⁴/ + /ɚ²/ → [pɚ⁴] 'very' (as compared to [peɪ⁴] 'double'). In closed syllables, if the coda is /n̪/, the vowel is rhotacized, while the coda is deleted, e.g., /pan̪⁴/ + /ɚ²/ → [pa˞ ⁴] 'partner' (as compared to [pan̪⁴] 'companion'); if the coda is /ŋ/, the vowel is nasalized and rhotacized, e.g., /xwɑŋ²/ + /ɚ²/ → [xwɑ̃˞²] 'yolk' (as compared to [xwɑŋ²] 'yellow').

Syllable Structure & 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 /jɛ¹/ 'duck', /w/ in /wɛ¹/ 'frog' and /ɥ/ in /ɥe¹/ 'to restrict'. Onsetless syllables are also possible, as in /aɛ¹/ 'sad'. However, in the old variety Tianjin Mandarin as reported in Han (1993b), onsetless syllables are not allowed before the rhymes of /an̪/, /ɑŋ/, /ən̪/, /ɤ/, /aɛ/, /ɑɔ/, or /əʊ/; a nasal onset /n̪/ is obligatory. Coda is optional, and only /n̪/ and /ŋ/ are allowed (as in /t̪an̪¹/ 'single' or /t̪ɑŋ¹/ 'when'). /n̪/ is the only consonant in the language that can appear both at the beginning and end of a syllable (as in /n̪ɛ⁴/ 'to include' and /t̪an̪¹/ 'single').

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

1) High front vowels /i, y, ɪ, ʏ/ and the corresponding glides /j, ɥ/ cannot follow dentoalveolar obstruents /ts̪, tsʰ̪, s̪/, postalveolar /tʃ̪, tʃʰ̪, ʃ̪/, or velar consonants /k, kʰ, x/.

However, /y/ and /ɥ/ can follow dentoalveolar /n̪, l̪/ (as in /n̪y^{3}/ ‘female’, /l̪y^{2}/ ‘donkey’, /n̪ɥe^{4}/ ‘to abuse’, /l̪ɥe^{4}/ ‘to omit’).

2) Alveolo-palatals /tɕ, tɕ^{h}, ɕ/ can only occur before high front vowels /i, y, ɪ, ʏ/ (as in /tɕi^{1}/ ‘chicken’, /tɕ^{h}i^{1}/ ‘seven’, /ɕi^{1}/ ‘west’, /tɕy^{1}/ ‘to live’, /tɕ^{h}y^{1}/ ‘maggot’, /ɕy^{1}/ ‘needs’, /tɕɪn̪^{1}/ ‘gold’, /tɕ^{h}ɪn̪^{1}/ ‘to invade’, /ɕɪn̪^{1}/ ‘heart’, /tɕʏn̪^{1}/ ‘army’, /tɕ^{h}ʏn̪^{2}/ ‘skirt’, /ɕʏn̪^{1}/ ‘to fumigate’), or their corresponding glides /j, ɥ/ (as in /tɕje^{1}/ ‘to connect’, /tɕ^{h}je^{1}/ ‘to cut’, /ɕje^{1}/ ‘to rest’, /tɕɥe^{2}/ ‘to feel’, /tɕ^{h}ɥe^{1}/ ‘to lack’, /ɕɥe^{1}/ ‘boots’). Given this restriction, the phonemic status of the alveolo-palatals has been a matter of debate, as relative to the dentoalveolar /ts̪, ts̪^{h}, s̪/, the postalveolar /tʃ̪, tʃ̪^{h}, ʃ̪/, and the velar /k, k^{h}, 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.

3) Mid-high vowels /e, o/ and mid-low vowel /ɛ/ have to co-occur with glides /j, w/, respectively, as in /t̪je^{1}/ ‘dad’, /t̪wo^{1}/ ‘more’, /t̪jɛn̪^{1}/ ‘bump’. Mid-high back vowel /ɤ/ cannot follow labial consonants /p, p^{h}, m, f/ in the younger-variety Tianjin Mandarin; the structure of /p, p^{h}, m, f/ + /ɤ/ has been only reported for the old variety in Han (1993b).

4) Glide /w/ cannot be followed by front vowels while /ɥ/ can only precede front vowels (as in /ɥe^{1}/ ‘to restrict’). /j/, however, can be followed by front, central and back vowels as in /ji^{1}/ ‘one’, /jɐ^{1}/ ‘duck’, /jɑŋ^{1}/ ‘central’.

5) Han (1993b) noted that in the old-variety Tianjin Mandarin, /n̪/ or /l̪/ do not co-occur with the rounded high front vowel /y/ in a number of open syllables. So, words like /n̪y^{3}/ ‘female’ and /l̪y^{4}/ ‘green’ in the young-variety Tianjin Mandarin would be pronounced as /n̪weɪ^{3}/ and /l̪weɪ^{4}/ in the old-variety Tianjin Mandarin.[1]

[1] Syllable structures as in /n̪weɪ^{3}/ and /l̪weɪ^{4}/ cannot be observed in the young-variety Tianjin Mandarin. Sound files for these two syllables are thus not provided.

Lexical Tones

	PHONETIC	ORTHOGRAPHIC	GLOSS
Tone 1	l̥ɑɔ^{1}	捞	‘to dredge up’
Tone 2	l̥ɑɔ^{2}	劳	‘hard-working’
Tone 3	l̥ɑɔ^{3}	老	‘old’
Tone 4	l̥ɑɔ^{4}	涝	‘to flood’

There are four full lexical tones in Tianjin Mandarin. Figure 9 shows the f0 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 f0 values were normalized so that f0 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 (1920)’s lexical tone annotation system, respectively.

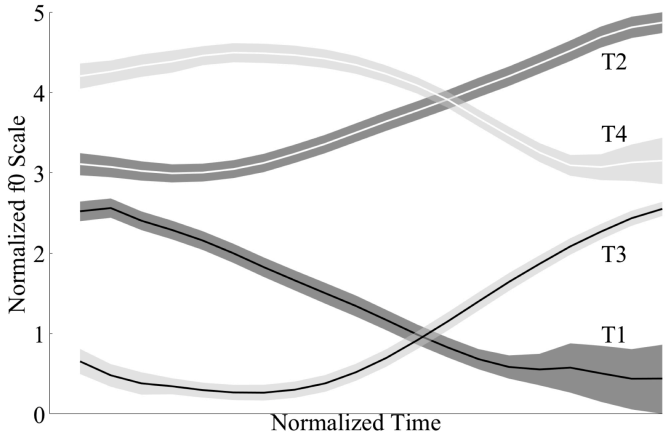


Figure 9 *Lexical tones in isolation. Lines stand for the mean. Gray areas stand for ±1 standard error of mean. Tone 1 (T1) is illustrated with black line and dark gray area; Tone 2 (T2) with white line and dark gray area; Tone 3 (T3) with black line and light gray area; Tone 4 (T4) with white line and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

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 /l̥ɑɔ^{1}/ ‘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 /l̪ɑɔ^{2}/ '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 /l̪ɑɔ^{3}/ 'old'. Tone 4 (T4) is a high-falling tone which falls from the upper end to the mid of the pitch range, as in /l̪ɑɔ^{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 f0 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).

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

T1	T2	T3	T4	
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

Tonal Variability

When lexical tones are produced in connected speech, their f0 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 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 details on tonal coarticulation in Tianjin Mandarin, readers are referred to Zhang & Liu (2011) and Li & Chen (2016).

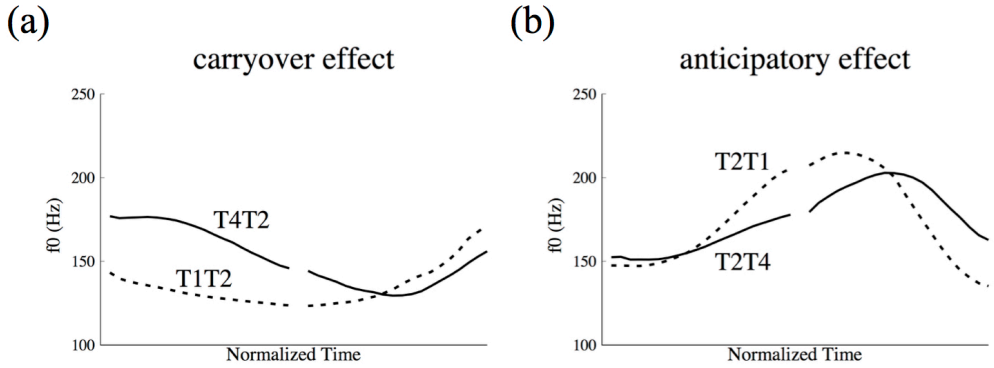


Figure 10 *f0 realization of T2 when connected with a high tone (T4) vs. a low tone (T1). T2 as the second tone in (a), T2 as the first tone in (b). Time is normalized.*

As shown in Figure 10(a), the f0 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^{4} məʊ^{2}/ 'stratagem'), the onset f0 realization of the second T2 is clearly higher than that following a low tone such as T1 (as in /kweɪ^{1} mwo^{2}/ 'scale'). Similar carryover effects could be observed from the comparison of T4T4 (as in /n̪eɪ^{4} mu^{4}/ 'inside story') vs. T1T4 (as in /kweɪ^{1} mi^{4}/ '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 /paɛ^{2}

mɑɔ[1]/ ‘white cat’), the offset f0 realization of the first T2 shows faster rate of f0 rise than that before a high tone such as T4 (as in /tʃ̻ʰəʊ[2] mi[4]/ ‘dense’). Similar anticipatory effects could also be observed from the comparison of T4T2 (as in /tɕi[4] məʊ[2]/ ‘stratagem’) vs. T4T3 (as in /mi[4] mɐ[3]/ ‘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 f0 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).

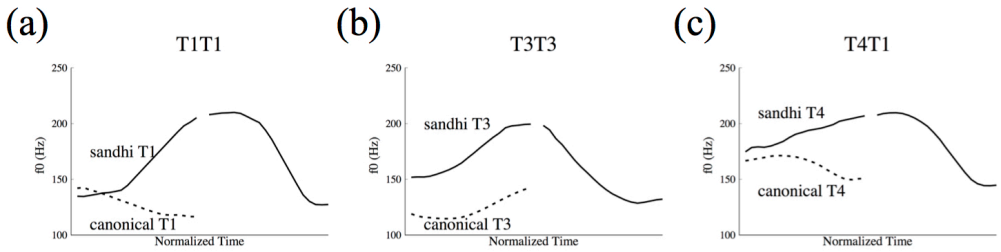


Figure 11 *f0 realization of three tone sandhi patterns in Tianjin Mandarin. Solid lines for the f0 realization of the sandhi patterns, dashed lines for the canonical f0 realization. Time is normalized.*

It can be seen from Figure 11 that in all the three tonal combinations, the first tone is realized with a drastically different f0 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 f0 contour any more as its canonical form (dashed line in Figure 11(a)). Instead, the f0 offset of the sandhi tone (solid line in Figure 11(a)) is raised to a great extent (as in /tɕjɐ[1] mɑɔ[1]/ ‘domestic cat’). In T3T3, the first T3 is realized with a high-rising f0 contour (solid line in Figure 11(b)), which is different from the low-dipping/low-rising f0 contour of

the canonical T3 (dashed line in Figure 11(b)), as in /wu^{3} n̪y^{3}/ ‘dancing girl’. In T4T1, the first tone shows a high-rising f0 (solid line in Figure 11(c)), as in /xəʊ^{4} mɐ^{1}/ ‘stepmother’. It is again very different from the high-falling f0 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.

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 /t̪^{h}wo^{2} n̪jɑɔ^{3}/ ‘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 /n̪aɛ^{3} ma^{1}/ ‘nanny’) and T2T1 (dashed line in Figure 12(b); as in /paɛ^{2} mɑɔ^{1}/ ‘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).

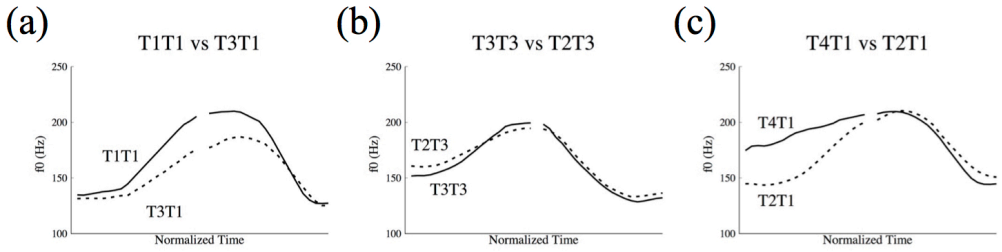


Figure 12 *f0 realization of three tone sandhi patterns in Tianjin Mandarin compared to their purported sandhi output tone patterns. Solid lines for the f0 realization of the sandhi patterns, dashed lines for the purported sandhi output tone patterns. Time is normalized.*

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^{3} pɐ^{3} tɕje^{2}/ 'torch festival') and T2T3T3 (Figure 13(b); as in /xwɐ^{2} pjɑɔ^{3} tɕjɑŋ^{3}/ 'Marble Pillar Award'), the first T3 is realized with a rising f0 contour (comparable to Figure 11(b)), indicating the application of tone sandhi in both contexts.

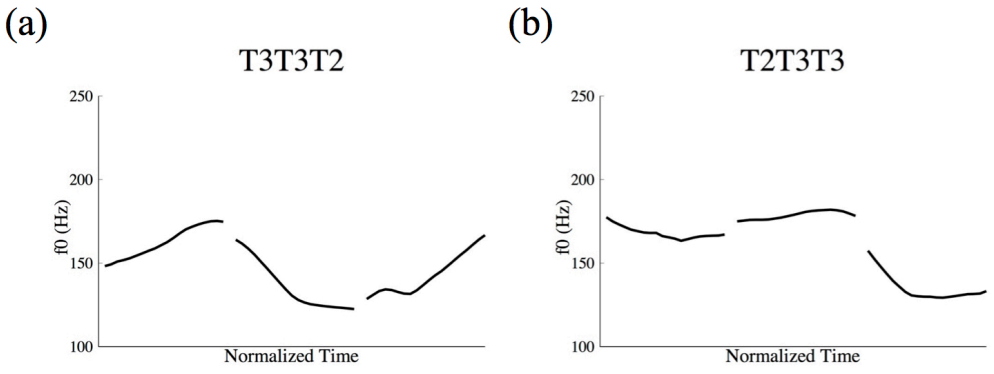


Figure 13 *f0 realization of trisyllabic sequences when T3T3 is left-aligned (a. T3T3T2) vs right-aligned (b. T2T3T3). Time is normalized.*

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^{2} t̪əŋ^{1} tɕ^{h}y^{1}/ '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^{1} ʃɑŋ^{1} tɕy^{2}/ 'Trade and Industry Bureau', tone sandhi does not apply since given 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 /tʂɤ^{2} tɕi^{4} pan^{1}/ 'acrobatic team'), while tone sandhi does not apply when the sequence is left-aligned as in T4T1T2 (Figure 15(a); as in /tʰaɛ^{4} kʰuŋ^{1} tʃʰwɐn^{2}/ 'spaceship'). For more experimental data on trisyllabic tone sandhi in Tianjin Mandarin, see Li & Chen (2016).

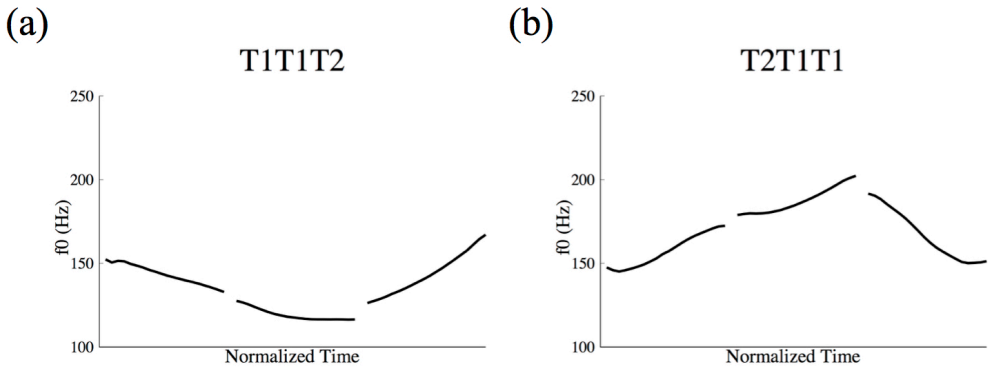


Figure 14 *f0 realization of trisyllabic sequences when T1T1 is left-aligned (a. T1T1T2) vs right-aligned (b. T2T1T1). Time is normalized.*

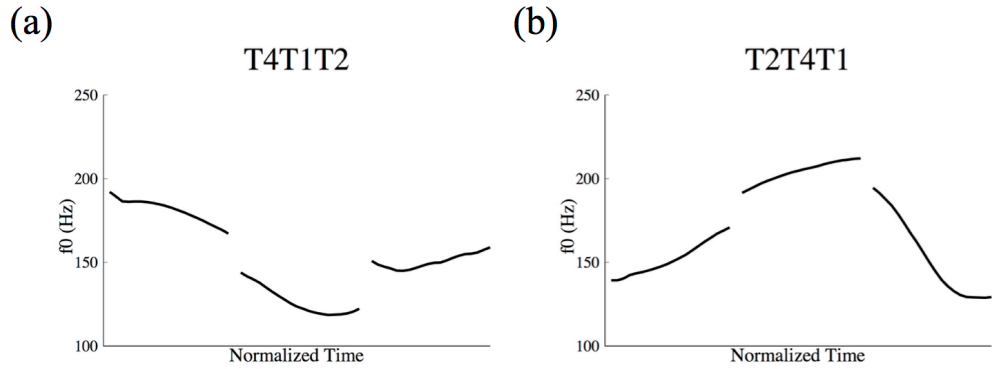


Figure 15 *f0 realization of trisyllabic sequences when T4T1 is left-aligned (a. T4T1T2) vs right-aligned (b. T2T4T1). Time is normalized.*

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 /kɤ^{1} kɤ/ '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 [kɤ^{1} 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 /t̬ə/ in /wo^{3} t̬ə/ 'mine'), lexical items (e.g., /l̬i/ in /pwo^{1} l̬i/ 'glass'), diminutive words (e.g., /kɤ/ in /kɤ^{1} kɤ/ 'elder brother'), and reduplication (e.g., /k^{h}an̬/ in /k^{h}an̬^{4} k^{h}an̬/ '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 /peɪ^{1} t̬ʃə mɐ^{1}/ '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 6 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).

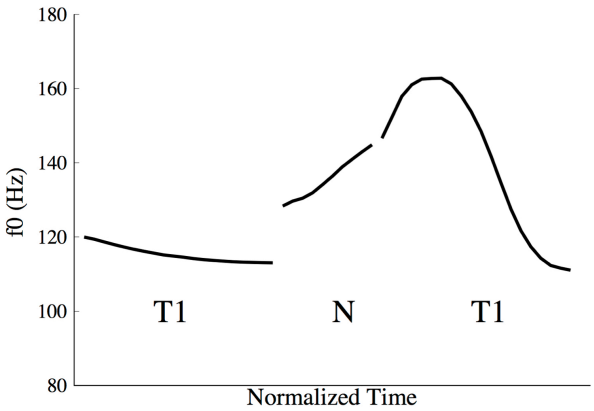


Figure 16 *f0 realization of one neutral tone (N) embedded between two T1s. Time is normalized.*

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 as 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 /tʰɐ¹ wo¹ mɐ¹ mɐ mən̰ t̰ə mɑɔ¹ t̰ʰi¹ l̰wan̰⁴ l̰ə n̰ɐ⁴ kə ɕjen̰⁴ tɕʰjəʊ²/ ‘He said mothers’ cats messed up that cotton ball’. The raised f0 realization of neutral tone could 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 /t̰ʰɐ¹ ʃ̰wo¹ mɐ¹ mɐ mən̰ t̰ə tʂ̰əŋ¹ tɕjɐ¹ lə ʂan̰¹ paɛ³ kʰwaɛ⁴ tɕʰjen̰²/ ‘He said mothers’ had increased by 300 yuan’.

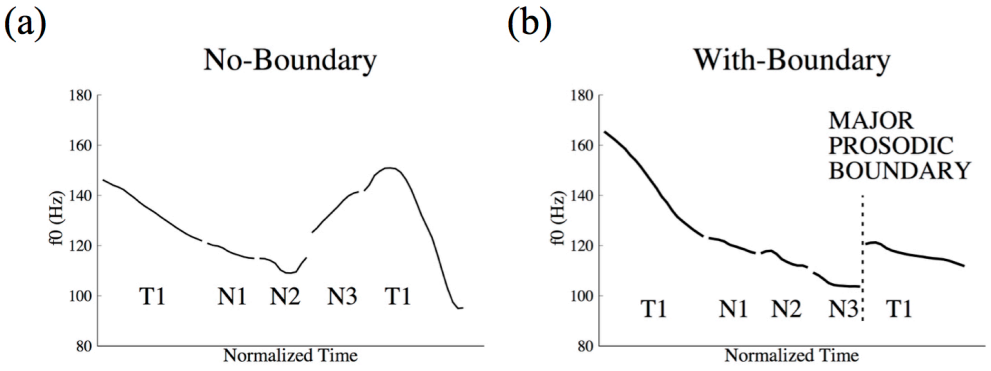


Figure 17 *f0 realization of three neutral tones (N1N2N3) embedded between two T1s without (a) or with (b) a major prosodic boundary following the neutral tone sequence. Time is normalized.*

Transcription of recorded passage ‘North Wind and the Sun’

This 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., ⁽¹⁾ 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.

jəʊ^{3} ji$^{(4)}$ ʈʰjɛɳ^{1} | peɪ^{3} fəŋ^{1} xɤ^{2} ʈʰaɛ^{4} jɑŋ | tʃəŋ^{4} tʂaɛ^{4} ʃwo^{1} | ʃeɪ^{2} ʈə | pəɳ^{3} ʃɻ̩ ʈɐ^{4} || tʃəŋ^{4} xɑɔ^{3} | jəʊ^{3} kə | tʃʰwaɳ^{1} ʈəʊ^{3} pʰəŋ ʈə ɻəɳ^{2} | tʂəʊ^{3} kwo^{4} ɭaɛ || ʈʰɐ^{1} ʎjɐ^{3} ʃwo^{1} | ʃeɪ^{2} ɳəŋ^{2} ɻɑŋ^{4} | ɳeɪ^{4} kə ɻəɳ^{2} | ʈʰwo^{1} ʈjɑɔ | ʈəʊ^{3} pʰəŋ | tɕjəʊ^{4} ʂwaɳ^{4} | ʃeɪ^{2} ɭi^{4} xaɛ || ɻaɳ^{2} xəʊ^{4} | peɪ^{3} fəŋ^{1} tɕjəʊ^{4} | pʰɪɳ^{1} ɭə mɪŋ^{4} ʈə tʃʰweɪ^{1} || ʈaɳ^{4} ʃɻ̩ | ʈʰɐ^{1} tʃʰweɪ^{1} ʈə | ɥe^{4} ɭi^{4} xaɛ | ɳeɪ^{4} kə ɻəɳ^{2} | tɕjəʊ^{4} pɐ$^{(3)}$ ʈəʊ^{3} pʰəŋ | kwo^{3} ʈə ɥe^{4} tɕɪɳ^{3} || peɪ^{3} fəŋ^{1} | meɪ^{2} tʃɤ^{2} | tʃɻ̩$^{(3)}$ xɑɔ^{3} | fɑŋ^{4} tɕʰi^{4} || tɕje^{1} tʃə | ʈʰaɛ^{4} jɑŋ | ʃaɛ^{4} ʈə | ɻɤ^{4} xʊŋ$^{(1)}$ xʊŋ^{1} ʈə | ɳeɪ^{4} kə ɻəɳ^{2} | ɭi^{4} mɐ$^{(3)}$ pɐ^{3} | ʈəʊ^{3} pʰəŋ | ʈʰwo^{1} ʈjɑɔ^{4} ɭə || ɻaɳ^{2} xəʊ^{4} | peɪ^{3} fəŋ^{1} | tʃɻ̩^{3} ɳəŋ^{2} | ɻəɳ$^{(4)}$ ʃu^{1} ||

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。

Acknowledgments

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 editor 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 & Weenink, David. 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: Harvard University Press.

Chen, Matthew Y. 2000. *Tone sandhi: Patterns across Chinese dialects.* Cambridge: Cambridge University Press.

Chen, Yiya. 2015. Qingsheng (Neutral tone). In Rint Sybesma (ed.), *Encyclopedia of Chinese language and linguistics*. Leiden: Brill. http://referenceworks.brillonline.com/browse/encyclopedia-of-chinese-language-and-linguistics/ (accessed 18 Jan 2017).

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 (ed.). 1993a. *Tianjin fangyan* [Tianjin Mandarin]. Beijing: Yanshan Chubanshe.

Han, Gendong. 1993b. Tianjinren zenyang xue putonghua [How do Tianjiners learn Standard Chinese]. In Gendong Han (ed.), *Tianjin fangyan* [Tianjin Mandarin]. Beijing: Yanshan Chubanshe.

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. Etúdes sur la phonologie Chinoise [Studies of Chinese phonology]. In *Archives d'Etúdes Orientales.* Leiden & Stockholm: E. J. Brill.

Ladefoged, Peter. 1999. American English. In *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet*, 41–44. International Phonetic Association.

Ladefoged, Peter. 2006. *A course in phonetics*. London: Wadsworth Publishing.

Ladefoged, Peter & Ian Maddieson. 1996. *The sounds of the world's languages*. Oxford: Blackwell Publishers.

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 C.-T. James Huang, Y.-H. Audrey Li, & Andrew Simpson (eds.), *The handbook of Chinese linguistics*, 369–399. Wiley Blackwell.

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 C.-T. James Huang, Y.-H. Audrey Li, & Andrew Simpson (eds.), *The Handbook of Chinese Linguistics*, 400–421. Wiley Blackwell.

Liu, Sixun. 1993. Tianjin fangyan de yuyin biandiao [Tone sandhi in Tianjin Mandarin]. In Gendong Han (ed.), *Tianjin fangyan* [Tianjin Mandarin]. Beijing: Yanshan Chubanshe.

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 cobination 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 and Zhongwei Shen (eds.), *'Lezaiqizhong – Wangshiyuan jiaoshou qishi huadan qingzhu wenji*, 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 Manadrin]. *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]. In *Qinzhu zhongguo shehui kexueyuan yuyan yanjiusuo jiansuo sishiwu zhounian xueshu lunwenji*. 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 & Liu, Jiang. 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.