

The Taiwanese Accent of Mandarin: A Corpus Analysis of Marking Influences on Accented L2

Yuchau E. Hsiao (蕭宇超)

Distinguished Professor

Graduate Linguistics Institute

National Chengchi University, Taipei 11605, Taiwan

ychsiao@nccu.edu.tw

1. Introduction

- The Taiwanese people who are born in the 1940s or earlier are known as “old Taiwanese,” and many of them speak Mandarin with a heavy accent.
- Taiwanese accented Mandarin, hereafter, T-Mandarin.
- 2. The phonotactic basics of Taiwanese and Mandarin. 3. The Corpus. 4. The patterns. 5. Summary. 6. Theoretical Generalization.

2. Phonotactic Basics

(1) Taiwanese Phonetic Vocalic List

Vowels		Glides	
i ĭ		u ũ	j w
e ě		(ɣ)o ɔ̃	
ɛ ě	ə	ɔ̃	
	a ǎ		

(2) Mandarin Phonetic Vocalic List

Vowels		Glides	
i y	ɿ ʅ	u	j ɥ w
e		ɤ o	
ɛ	ə	ɔ	
a(æ)		ɑ	

- Taiwanese and Mandarin have similar oral vowels and glides, except that the high front rounded vowel [y] and glide [ɥ] are absent in Taiwanese.
- In Taiwanese, most vowels have nasal counterparts, while nasal vowels in Mandarin occur before nasal codas.
- In certain sub-dialects of Mandarin, [æ] is an allophone of [a]. However, I have observed that [æ] is not heard among the Mandarin speakers in Taiwan.
- There are four vocalic sequences in Mandarin that are not found in Taiwanese, including [je], [ei], [wo], and [ou].
- Two apical vowels in Mandarin: [ɿ] after dental obstruents and [ʅ] after retroflex consonants; these vowels are produced with both the tongue body and the tongue tip.
- I propose that the apical vowels are [+back] in nature, due to two reasons. First, the apical vowels and the back high vowel cannot occur after the alveo-palatals, [tɕ], [tɕʰ] and [ɕ], which are restricted before the front high vowels, [i] and [y]. Second, the production of the apical vowels involves the raise of the tongue back.

(3) Taiwanese Phonetic Consonantal List

p p ^h	t t ^h	k k ^h	ʔ
b	d/l	g	
	ts ts ^h	tɕ tɕ ^h	
	s z	ɕ ʒ	h
m	n	ŋ	

(4) Mandarin Phonetic Consonantal List

p p ^h	t t ^h	k k ^h	ʔ
	ts ts ^h	tɕ tɕ ^h	tɕ tɕ ^h
	f s	ɕ ʒ	tɕ x
m	n		ŋ
	l	r	

- Phonemically, Mandarin has no voiced stop, though phonetically a voiceless stop might be voiced in an intervocalic position. However, this voicing pattern is not found among the Mandarin speakers in Taiwan.
- Taiwanese allows three voiced stops in the surface, [b], [d], and [g]. The closure in the production of [d] in Taiwanese is very slack such that the pronunciation of [d] is very close to [l], especially before a low vowel; these two sounds are usually considered the same segment in this dialect.
- Mandarin has two voiceless fricatives that are not found in Taiwanese, namely, the labiodental [f] and the velar [x]. However, the velar [x] is usually pronounced further back before back vowels and perceptually sounds like [h].
- There are five retroflex consonants in Mandarin, including four retroflex obstruents, [tʂ], [tʂʰ], [ʂ] and [ʐ], and a retroflex liquid, [r].
- The voiced stops (D) in Taiwanese are onsets, in complementary distribution with the nasal onsets (N), [m], [n], and [ŋ]; namely, the output sanctions DV and N \tilde{V} , but not D \tilde{V} nor NV. On the other hand, Mandarin oral and nasal onsets indicate no such constraint.
- In Taiwanese, a nasal coda cannot occur after a nasal vowel, but in Mandarin a nasal coda nasalizes its preceding vowel.

3. The Corpus

- Recording: 8 hours, 1260 words collected.
- Taiwanese speakers: 2 females + 2 males, aged 71~73, of primary education or below.

(5) Exclusive Patterns of Mandarin

- The apical vowels, [ɿ] and [ʉ].
- The high front rounded vowel, [y], and glide, [ɥ].
- The vocalic sequences, [ei], [je], [ou] and [wo].
- The nasal syllables, N \tilde{V} N and C \tilde{V} N
- The retroflex consonants, [tʂ], [tʂʰ], [ʂ], [ʐ] and [r].
- The voiceless labiodental fricative [f].

(6) Three kinds of pronunciations

<u>M-Taiwanese</u>	<u>Taiwanese</u>	<u>Mandarin</u>	<u>Gloss</u>
sa	ʂã	swa	'sand'

- Most Taiwanese words have two readings, vernacular and literary. Literary pronunciation is not considered here, since the informants merely received primary education.

(7) Phonological processes examined in T-Mandarin

- Will the apical vowels be interpreted as other vowels?
 - Will [y] and [ɥ] be derounded?
 - Will [ei], [je], [ou] and [wo] be monophthongized?
 - Will the vowels in N \tilde{V} N and C \tilde{V} N be denasalized?
 - Will [tʂ], [tʂʰ], [ʂ], [ʐ] and [r] be deretroflexed?
 - Will [f] be interpreted as other fricatives?
- The rounding of the apical vowels and the derounding of the interior vowel and glide will introduce their less marked, peripheral counterparts. The delabialization of [f] will obtain a less marked fricative. Monophthongization, vowel denasalization and deretroflexion are also

operations that show a decrease in markedness.

4. The Patterns

4.1 Rounding

- The corpus has 108 words that contain the apical vowels, [ɿ] and [ʮ], and all of them are rounded as [ʊ].

(8) Rounding

<u>T-Mandarin</u>	<u>Mandarin</u>	<u>Taiwanese</u>	<u>Gloss</u>
tsu	tʂʊ	tswaʔ	‘paper’
ts ^h u	tʂ ^h ʊ	tjaʔ	‘eat’
su	ʂʊ	tsap	‘ten’
zu	ʐʊ	zit/lit	‘day’
tsu	tsɿ	tsu	‘fund’
ts ^h u	ts ^h ɿ	te ^h i	‘sting’
su	sɿ	ei/si	‘four’

4.2 Derounding

- There are 126 words with [y]/[ʮ] in the corpus, and 122 of them are derounded as [i]/[j]. That is, 96.8% of the data display the Taiwanese accent, while only 3.2% of them are pronounced without an accent.

(9) Derounding set 1: identical vowels/glides

<u>T-Mandarin</u>	<u>Mandarin</u>	<u>Taiwanese</u>	<u>Gloss</u>
ei	ɛy	hi	‘swear’
te ^h i	te ^h y	k ^h i	‘go’
ni	ny	lu/li	‘female’
ji	ɲy/ʎy	hi	‘fish’
li	ly	li	‘donkey’
tje	tje	tjok	‘peerage’

(10) Derounding set 2: different vowels/glides

<u>T-Mandarin</u>	<u>Mandarin</u>	<u>Taiwanese</u>	<u>Gloss</u>
ji	ɲy/ʎy	gu	‘encounter’
te ^h i	te ^h y	k ^h u	‘district’
li	ly	lut	‘rate’
eje	eje	hak	‘learning’
tjen	tjɛn	kwan	‘donate’
njɛ	nɲɛ	gjət	‘harsh’

- One question thus arises as to whether the derounding in T-Mandarin simply results from the cognate correlation instead of vowel markedness. The answer can be pursued from (10), where

the vowels in T-Mandarin are not the same as those of the Taiwanese pronunciations, but the Mandarin vowels are derounded in spite of everything. This fact shows that the accent is not as simple as cognate correlation, but rather, it occurs by replacing a marked form with an unmarked one.

4.3 Monophthongization

- In the corpus, 190 out of 199 [ei] sequences are pronounced as [e], found in 95.4% of the data, and 134 out of 166 [je] sequences are pronounced as the monophthong, found in 80.7% of the data; similarly, 203 out of 210 [ou] sequences are pronounced as [o], found in 96.6% of the data, and 160 out of 189 [wo] sequences are pronounced as the monophthong, found in 84.6% of the data.

(11) Monophthongization

<u>T-Mandarin</u>	<u>Mandarin</u>	<u>Taiwanese</u>	<u>Gloss</u>
p ^h e	p ^h ei	pwe	‘to accompany’
tse	teje	tejo?	‘to borrow’
ko	kwo	kok	‘nation’
t ^h o	t ^h ou	t ^h au	‘to steal’

4.4 Denasalization

- There are 78 syllables in the corpus that consist of C[~]VN structures, but all of them are pronounced as CVN. There are 93 syllables that consist of N[~]VN structures: 89 of them are pronounced as CVN, constituting 95.7%, and only 4 of them are pronounced as N[~]VN, constituting 4.3%.

(12) Denasalization

	<u>T-Mandarin</u>	<u>Mandarin</u>	<u>Taiwanese</u>	<u>Gloss</u>
a.	loŋ	lõŋ	ljəŋ	‘dragon’
	ein	ẽin	sin	‘new’
b.	ben	mjẽn	mi □	‘noodle’
	ljaŋ	njãŋ	njũ	‘mother’

4.5 Deretroflex

- The corpus has 246 words that contain retroflex onsets, and 230 of them are deretroflexed, found in 93.4% of the data. There are 42 words that have the retroflex coda, and all of them undergo coda drop.

(13) Deretroflex

<u>T-Mandarin</u>	<u>Mandarin</u>	<u>Taiwanese</u>	<u>Gloss</u>
tsu	tʂu	ti	‘pig’
ts ^h a	tʂ ^h a	te	‘tea’
sa	ʂa	swa	‘sand’
zu, lu	zu	zip/lip	‘enter’

ʔə ʔər zi/li ‘son’

4.6 Reconfiguration

- In the corpus, 50 syllables have [f] onsets; 32 of them are pronounced as [hw], constituting 64%, and 18 of them are pronounced as [h], constituting 36%.

(14) Set 1: [f] reconfigured as [hw]

<u>T-Mandarin</u>	<u>Mandarin</u>	<u>Taiwanese</u>	<u>Gloss</u>
hwe	fei	pwi	‘fat’
hwaŋ	fǎŋ	paŋ	‘release’
hwa	fa	hwat	‘punish’

(15) Set 2: [f] delabialized as [h]

<u>T-Mandarin</u>	<u>Mandarin</u>	<u>Taiwanese</u>	<u>Gloss</u>
ho	fwo	hut	‘Buddha’
ho	fou	ho □	‘not’
hoŋ	fuŋ	p ^h aŋ	‘seam’

5. Summary

(16) Vocalic Markedness

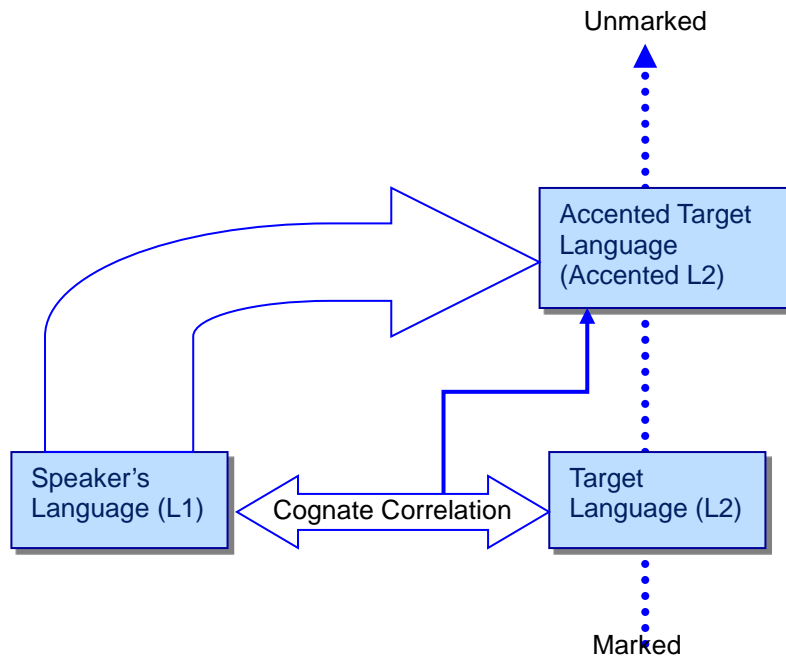
- Interior vowels are marked, and tend to be replaced by peripheral vowels in the accent; [y] and [ɥ] are respectively derounded as [i] and [j] in T-Mandarin.
- Apical vowels are marked, and tend to be interpreted as non-apical in the accent; [ɪ] and [ɯ] are rounded as [u] in T-Mandarin.
- Diphthongs (and the glide-vowel sequences) are marked, and tend to be replaced by monophthongs in the accent; in T-Mandarin, [ei] and [je] are simplified as [e], and [ou] and [wo] as [o].
- Nasal vowels are marked, and tend to be replaced by oral vowels in the accent; the nasal vowels are denasalized before nasal codas in T-Mandarin.

(17) Consonantal Markedness

- Retroflexed consonants are marked, and tend to be replaced by nonretroflexed consonants or omitted in the accent; the onsets, [tʂ], [tʂ^h], [ʂ] and [ʐ] are deretroflexed as [ts], [ts^h], [s] and [z] in T-Mandarin, and the coda, [r], is deleted.
- The combination of labiality and voiceless features is marked, and the features tend to be split or delabialized in the accent; [f] is reconfigured as [hw] or delabialized as [h] in T-Mandarin.
- The feature [dorsal] is more marked than [pharyngeal], and the former tends to be replaced by the latter in the accent; [x] is devoiceless as [h] in T-Mandarin.

6. Theoretical Generalization

(18) A Marking-based Model of Accent Formation



- In this schema, the target language (L2) can be a local dialect or a foreign language. The speaker's language (L1) is the native language of the speaker. The accented target language (accented L2) is the default language used by the speaker. The vertical dotted arrow represents the path of universal marking, and the right-headed hollow arrow indicates the imposition of the L1 elements on the accented L2. The horizontal bidirectional hollow arrow denotes the cognate correlation between L1 and L2. The elbow type arrow represents these cognate effects on the accented L2.
- Phonological influences in language contact inevitably involve systematic adjustments of segments. Such adjustments often result in the emergence of the universally unmarked.
- The output of an accent is by and large the intersection between the target language (L2) and the speaker's language (L1). It is the omission of marked forms that causes the accents.
- The marked forms of L2 that are not present in L1 will be replaced by the unmarked forms of L1. Conversely, if the exclusive forms of L2 are unmarked, they will not be replaced by the marked forms of L1, but will be retained in the accented L2.
- Taiwanese and Mandarin share a substantial amount of cognates. Some Mandarin words containing [f] correspond to those containing [h] in Taiwanese. However, this cognate correspondence observes the marking relation of segments in accent formation.

Acknowledgments

I would like to thank the Ministry of Science and Technology in Taiwan for sponsoring this trip.

Source

Hsiao, Yuchau E. (蕭宇超). 2011. Universal marking in accent formation: evidence from Taiwanese-Mandarin and Mandarin-Taiwanese. *Lingua* 121: 1485–1517.
[doi:10.1016/j.lingua.2011.04.002](https://doi.org/10.1016/j.lingua.2011.04.002)