Segmental and Tonal Errors in L2 Mandarin Speech Produced by Australian English Learners

Wentao Gu [1, 2], *Ting Zhang* [1], *Chiharu Tsurutani* [3]

[1] Research Center for Language Information Technologies, Nanjing Normal University, China
[2] Graduate School of Information Science and Technology, The University of Tokyo, Japan
[3] School of Languages and Linguistics, Griffith University, Australia
wtgu@njnu.edu.cn, tingtina_zhang@126.com, c.tsurutani@griffith.edu.au

Abstract

Despite the increasing importance of learning Mandarin in international communities, the phonetics of L2 Mandarin has been studied far less than that of L2 English. In this study, L2 Mandarin speeches by the Australian beginner-level learners were collected on the basis of a phonetically balanced corpus. A systematic analysis of segmental and tonal errors showed that Australian learners made errors not only in tones but also in some initial consonants and finals which were absent in English. Most errors are associated with the L1 transfer effect, and some of them are also ascribed to the confusion caused by the Pinyin spelling. Our findings will be beneficial not only in the pedagogical sense – those frequently observed errors can be treated purposefully in language teaching, but also for the development of computer-assisted language learning programs.

Index Terms: L2 Mandarin, Australian English learner, tone, initial, final, L1 transfer

1. Introduction

It is known that L2 errors produced by the learners from a certain L1 background tend to have common patterns which reflect the transfer of their L1 phonology. While the error patterns in L2 English have been well documented, research into L2 Mandarin by native English speakers is far behind. Up until now, studies on L2 Mandarin have focused more on tones than on segments [1-4]. Systematic data collection for L2 Mandarin speech produced by native Australian speakers has not been reported, despite the increasing importance of learning Mandarin as a second language. In this study, we collected the Mandarin utterances produced by 33 Australian learners at the beginner level, and then analyzed the segmental and tonal errors in the data. This phonetic study on L2 errors will be useful, not only for language teaching, but also for developing computer assisted language learning programs.

2. Mandarin pronunciation

In terms of the conventional Chinese phonology, a syllable in Mandarin can be divided into an optional initial and a final which carries a tone. The initial is a consonant, while the final consists of an optional glide, a vowel nucleus (monophthong or diphthong), and an optional nasal coda. Mandarin has 21 initial consonants and 37 finals [5]. Also, Mandarin has five tones, including four lexical tones and a neutral tone. The initial consonants, finals and tones in Mandarin are listed in Tables 1-3 below.

In contrast, Australian English has 22 consonants, 12 vowels, and no tones. The frequent difficulties native English speakers have in producing Mandarin consonants, vowels and tones are summarized below, according to the literature [6-7] and the L2 Mandarin teaching practice of the second author.

Initial consonants: Native English speakers are not familiar with the following four sets of initial consonants in Mandarin, as highlighted in Table 1:

- Retroflex: “zh” [tʂ], “ch” [tʂʻ], “sh” [ʂʻ], r [ʐ].
- Alveolo-palatal: “j” [tɕ], “q” [tɕʻ], “x” [ɕ].
- Alveolar affricates: “z” [ts], “c” [tsʻ].
- Unaspirated voiceless stops: “b” [p], “d” [t], “g” [k].

It should be noted that in English the stops “b”, “d”, “g” are partially voiced, whereas in Mandarin, they are all voiceless.

Finals: The finals deemed difficult for native English speakers are highlighted in Table 2, where G, V and N indicate glide, vowel and nasal, respectively. In particular, the rounded vowel “ü” [y], the finals starting with the rounded glide “ü” [y], and the nasal finals are new to native English speakers. The glide in GVN tends to be deleted by native English speakers.

Tones: The detailed descriptions of Mandarin tones are listed in Table 3. Tone is usually the primary obstacle for L2 learners whose native languages are non-tone languages. For example, the confusion between T2 and T3 is frequently observed in L2 Mandarin learners.

Table 1. *Initial consonants in Mandarin.*

	Labial	Alveolar	Retroflex	Palatal	Velar
Stop	b [p]	d [t]			g [k]
	p [pʰ]	t [tʰ]			k [kʰ]
Affricate		z [ts]	zh [tʂ]	j [tɕ]	
		c [tsʰ]	ch [tʂʰ]	q [tɕʰ]	
Fricative	f [f]	s [s]	sh [ʂ]	x [ɕ]	h [x]
Nasal	m [m]	n [n]			
Liquid		l [l]	r [ʐ]		

Table 2. *Finals in Mandarin.*

(G)V		i [i]	u [u]	ü [y]
	a [a]	ia [ia]	ua [ua]	
	o [o]		uo [uo]	
	e [ɤ]	ie [iɛ]		üe [yɛ]
	i [ɿ]			
	i [ʅ]			
	er [ɚ]			
(G)VV	ai [æi]		uai [uæi]	
	ei [ei]		uei [uei]	
	ao [ɑu]	iao [iɑu]		
	ou [ou]	iou [iou]		
(G)VN	an [an]	ian [iɛn]	uan [uan]	üan [yɛn]
	en [ən]	in [in]	uen [un]	ün [yn]
	ang [ɑŋ]	iang [iɑŋ]	uang [uɑŋ]	
	eng [əŋ]	ing [iŋ]	ong [uŋ]	iong [yŋ]

Table 3. *Tones in Mandarin.*

ID	Name	Pitch value	Tone letter	Example
T1	High	55	˥	mā (mother)
T2	Rising	35	˧˥	má (hermp)
T3	Low	21(4)	˨˩	mǎ (horse)
T4	Falling	51	˥˩	mà (blame)
T0	Neutral	Context-dependent		ma (a question particle)

3. Speech data

3.1. Participants

Thirty three native speakers of Australian English (M=19, F=14) who were students of Chinese language courses at university in Brisbane participated in speech recording in return for a small payment. Their ages ranged from 18 to 63. At the time of recording, they had received 40~50 hours of classroom instruction at university. The recording was conducted from August to September 2013 at Griffith University and Queensland University of Technology, in Brisbane, Australia.

3.2. Materials

We prepared 20 short stimulus sentences composed of words and expressions familiar to L2 learners at the beginner level. These sentences, either declarative or interrogative, are phonetically balanced, including all the phonemes and tones that are deemed to be difficult for native English speakers. All sentences are listed below, showing Pinyin transcription, broad phonetic transcription, and English translation. At the time of recording, these sentences were presented both in Chinese characters and in Pinyin transcription.

1. Duìbuqǐ. Méiguānxi.
   [tuei4 pu0 tɕʰi3. mei2 kuan1 ɕɥ0]
   I am sorry. It doesn't matter.
2. Wǒ shì Àodàlìyà rén.
   [wo3 ʂɻ4 ɑu4ta4li4ja4 ʐən2]
   I am Australian.
3. Nǐ zuò shénme gōngzuò?
   [ni3 tsuo4 ʂən2mɤ kuŋ1tsuo4]
   What do you do?
4. Wǒmen yìqǐ qù hǎo ma? Hǎode.
   [wo3mən ji4 tɕʰi3 tɕʰy4 xɑu3 ma? xɑu3tə0]
   Could we go together? OK.
5. Wǒ lái jièshào yíxià, zhè shì WángBái.
   [wo3 læi2 tɕiɛ4ʂɑu4 ji2ɕia4, tʂɤ4 ʂɻ4 wɑŋ2pæi2]
   Let me introduce her. This is Wang Bai.
6. Wǒ xiànzài zài dàxué xuéxí Hànyǔ.
   [wo3 ɕiɛn4tsæi4 tsæi4 ta4ɕyɛ2 ɕyɛ2ɕi2 xan4jy3]
   I am studying Chinese at the university now.
7. Nǐ jiào shénme míngzi?
   [ni3 tɕiɑu4 ʂən2mɤ miŋ2tsɻ0]
   What's your name?
8. Wǒ cóng Yīngguó Lúndūn lái.
   [wo3 tsʰuŋ2 iŋ1kuo2 lun2tun1 læi2]
   I come from London.
9. Tā hěn xǐhuan chī jiǎozi.
   [tʰa1xən3 ɕɥ3xuan0 tʂʰʅ1 tɕiɑu3tsɻ0]
   She likes eating dumplings very much.
10. ChénDōng yǒudiǎnr lěng.
    [tʂʰən2tuŋ1 jou2tiɛn3ɚ ləŋ3]
    Chen Dong is feeling a little cold.
11. Rènshi nǐ wǒ yě hěn gāoxìng.
    [zʐən4ʂɻ0 ni3 wo3jɤ3 xən3 kɑu1ɕiŋ4]
    Nice to meet you, too.
12. Zhāng lǎoshī shì wǒmen de hànyǔ lǎoshī.
    [tʂɑŋ1 lɑu3ʂɻ1 ʂɻ4 wo3mən0 tɤ0 xan4jy3 lɑu3ʂɻ1]
    Teacher Zhang is our Chinese teacher.
13. Zhè běn shū hěn yǒu yìsi.
    [tʂɤ4 pən3 ʂu1 xən3 jou3ji4sɻ0]
    This book is very interesting.
14. Tā shì shuí, tā shì xuésheng ma?
    [tʰa1 ʂɻ4 ʂuei2, tʰa1 ʂɻ4 ɕyɛ2ʂəŋ ma0]
    Who is he? Is he a student?
15. Qǐng zài xiě yíbiàn cíyǔ.
    Please write the vocabulary once more.
    [tɕʰiŋ3 tsæi4 ɕiɛ3 ji2 piɛn4 tsʰɻ2 jy3]
16. ZhāngYuányuan xiǎng kàn diànshì.
    [tʂɑŋ1 juan2juan0 ɕiɑŋ3 kʰan4 tiɛn4 ʂɻ4]
    Zhang Yuanyuan wants to watch TV.
17. Wǒ zài yì jiā Rìběn gōngsī gōngzuò.
    [wo3 tsæi4 ji4 tɕia1 zɻ4pən3 kuŋ1sɻ1 kuŋ1tsuo4]
    I work in a Japanese company.
18. Wǒ xiànzài lái kéyǐ ma? Kéyǐ.
    [wo3 ɕiɛn4tsæi4 læi2 kʰɤ2 ji3 ma0? kʰɤ2 ji3]
    May I come now? Yes.
19. Bù, wǒ búshì Měiguó rén.
    [pu4, wo3 pu2ʂɻ4 mei3 kuo2 zʐən2]
    No, I am not American.
20. LínTóngtong shì tā de nǚpéngyou.
    [lin2 tʰuŋ2 tʰuŋ0 ʂɻ4 tʰa1 tɤ0 ny3pʰəŋ2 jou0]
    Lin Tongtong is his girlfriend.

Speech recording was conducted after the participants became sufficiently familiar with the materials and it was ensured that they knew the pronunciations. All participants were instructed to read each sentence aloud at their normal speech rates.

3.3. Transcription

Transcription was then conducted on the recorded speech using the Praat toolkit. Using the following guidelines, two layers of transcription were made by perceptual judgments:

1. On the basis of the text, Pinyin transcription for the nominal pronunciation.

(1) In the case of the tone *sandhi*, the *sandhi* tone (i.e., surface tone) is labelled instead of the underlying tone.

(2) In the case of tone neutralization, the neutral tone is labelled instead of the underlying tone.

2. Phonetic errors: Only the errors perceivable to native speakers are labelled; those phonetic variations commonly observed in native speakers' speech are not regarded as errors.

To ensure the reliability of transcription, we organized for two phonetically trained native Mandarin speakers to label the speech data independently following an iterative training session and discussion. After independent labelling, the Mandarin speakers crosschecked with each other until reaching a consensus on those inconsistent labels.

4. Results

The results for errors on all tones, initials and finals are summarized in Tables 4-6 and Figure 1. Table 4 shows the transition matrix for all tones; the highlighted items indicate the accuracy rates for each tone. The *sandhi* T3 (i.e., T3→T2/_T3) was counted separately. In comparison, the contour tones (i.e., T2 and T4) have higher error rates than the phonologically high or low tones (i.e., T1 and T3); the error rate is lowest for T1 and highest for T2. For native Australian speakers, T1 is not only the easiest tone to produce, but also was used frequently to replace other tones. Also, there is a fairly high mutual confusion between T2 and T3. In particular, the error rate for *sandhi* T3 is conspicuously lower than that for T2, suggesting that the production of *sandhi* T3 might be articulation-oriented instead of phonologically coded.

In Figure 1, the error rates for initial consonants were shown in descending order. There are no errors for [b, g, k, h, f, n, m]. As predicted, a large proportion of errors occurred on retroflexes, alveolar-palatals and alveolar affricates (i.e., "c" [tsʰ] and "z" [ts]). The detailed error patterns for the two alveolar affricates are listed in Table 5, which shows that L2 learners tend to replace a phoneme new to them by similar phonemes in their native language. In fact, the difficulty in producing alveolar affricates was also reported in studies of L2 Mandarin by native American English speakers [8].

As observed from Figure 1, most stops were produced correctly. It was reported that the speakers whose native language did not have a phonemic contrast between aspirated and unaspirated stops might substitute the voicing contrast for the aspiration contrast [9]. The aspiration contrast in Mandarin stops, however, is handled successfully by Australian learners, probably because the aspiration contrast in stops also exists in English. This was also reported in the study of L2 Mandarin in native German learners [10].

Table 4. *Transition matrix for tones (in %).*

Target \ Result	T0	T1	T2	T3	T4
T0	52.3	12.5	20.0	8.6	6.5
T1	1.0	66.1	13.4	12.4	7.2
T2	0	22.4	41.7	21.8	14.0
T2 (*sandhi* T3)	0	16.0	59.8	12.1	12.1
T3	0	11.8	18.4	61.8	8.0
T4	0.9	23.0	12.0	14.6	49.4

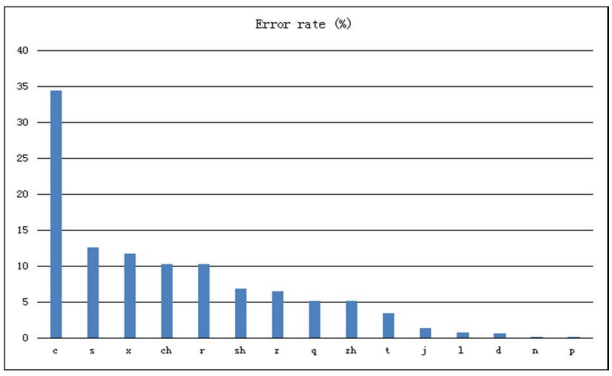

Figure 1: *Error rates for initial consonants (in %).*

Table 5. *Error distribution for affricates "c" [tsʰ] and "z" [ts].*

Target	c [tsʰ]	58	z [ts]	290
Result	ch [tʂʰ]	1	c [tsʰ]	10
	k [kʰ]	2	s [s]	6
	q [tɕʰ]	4	sh [ʂ]	1
	s [s]	5	[dz]	1
	sh [ʂ]	1	zh [tʂ]	1
	t [tʰ]	2		
	x [ɕ]	1		
	z [ts]	4		

Table 6. *Distribution and error rates for finals.*

Structure	Phoneme	Total number	Error rate (%)
V	ü [y]	29	82.8
	u [u]	203	12.8
	e [ɤ]	290	11.0
	(c)i [ɿ]	145	6.2
	i [i]	696	6.2
	(ch)i [ʅ]	348	2.3
	a [a]	377	1.6
	o [o]	348	0.6
GV VV GVV	üe [yɛ]	87	73.6
	uei [uei]	58	10.3
	ou [ou]	87	9.2
	ie [iɛ]	58	8.6
	ao [ɑu]	290	3.1
	ei [ei]	58	1.7
	ai [æi]	290	0.3
VN GVN	un [un]	58	29.3
	ang [ɑŋ]	116	6.9
	uan [uan]	116	4.3
	an [an]	377	2.7
	eng [əŋ]	116	2.6
	en [ən]	464	2.2
	ing [iŋ]	116	0.9
	ong [uŋ]	203	0.5

In Table 6, the error rates for finals are listed in descending order for each type of structure (the finals without errors were not listed). First of all, the rounded front vowel "ü" [y] has the highest error rate (82.8%), and 58.6% of the time it was mispronounced as [iou].

Secondly, the GV final "üe" [yɛ] shows a very high error rate (73.6%), and 65.5% of the time it was mispronounced as [uei]. The reason for this mispronunciation may be that [y] in this context shares the same Pinyin transcription "u" with [u].

Among nasal finals, "un" [un] was the most difficult one for Australian learners. In our stimulus sentences, "un" occurs only in syllables "lun" and "dun". Both "l" [l] and "d" [t] are alveolar consonants, for which the tongue tip slides rapidly from the back of the alveolar ridge to the back of the lower teeth [6]. During this process, Australian learners may fail to

round the lips. Thus, 20.7% of the time, “un” [un] was mispronounced as [ən]. This error might be triggered by the fact that the Pinyin spelling “un” is pronounced as [ən] in English.

In the previous study into the problems that native German learners [10] face in learning Mandarin, similar error patterns were also reported, such as:

- Substitution of T1 for T2~T4;
- Errors on retroflexes, alveolar-palatals and alveolar affricates.

However, some other segmental errors are more language-specific. For example, “ü” [y] and “e” [ɤ] are difficult for Australian learners but not so for German learners, as the German language has [y] and [ɤ] in its sound inventory.

5. Discussion

The pronunciation errors in L2 Mandarin speech have been investigated on the basis of 660 utterances, which, as far as we know, constitute the first L2 Mandarin speech database to be collected from native Australian learners. It has been found that both segmental and tonal errors occur mainly on the elements missing in the L1 (i.e., English), indicating the L1 transfer effect.

Tone is the major difficulty in L2 Mandarin produced by Australian learners whose native language is non-tonal. While the error rates for all tones are high, the contour tones (i.e., T2 and T4) are generally more difficult than the phonologically high or low tones (i.e., T1 and T3), suggesting that native English speakers are not skilled at producing big pitch change within a syllable. When the learners are not confident about the correct tones, it is often the case that they produce a high pitch, which sounds like T1.

There are also some difficulties in producing segments for L2 learners. The errors on certain initials (i.e., retroflexes, alveolar-palatals and alveolar affricates, especially, “c” [tsʰ]) and finals (e.g., “ü” [y], “üe” [yɛ], and “un” [un]) have been commonly observed – these sounds are generally not in the phonological inventory of Australian English. Some of these errors are also ascribed to the confusion between Pinyin system and English spelling. For example, “un” [un] tends to be mispronounced as [ən] like in English.

Our findings on L2 errors will be helpful for the pedagogical purpose. Those frequently observed errors should be treated more seriously in L2 teaching. For example, T2 and T4 should be given more intensive training than T1 and T3, whereas a set of contrastive training should be conducted purposefully between the confusable pairs such as “ü” [y] vs. “iou” [iou], “üe” [yɛ] vs. uei [uei], and “un” [un] vs. “en” [ən]. Moreover, the development of computer-assisted pronunciation assessment programs will also benefit from our findings by including those syllables prone to L2 errors as testing examples.

As the population of L2 Mandarin learners is rapidly increasing around the world, more efficient instructions based on empirical phonetic study will be needed, not only for L2 teaching/learning but also for the development of computer-assisted pronunciation assessment systems for L2 Mandarin. In this sense, the L2 error patterns deduced from our current speech data will assist further progress in this field.

6. Acknowledgements

This project was supported by the Griffith University research grant 2013, the National Social Science Foundation of China (10CYY009), and the Jiangsu Higher Education Institutions’ Key Research Base for Philosophy and Social Sciences (2010JDXM024).

7. References

[1] Wang, Y., Spence, M.M., Jongman, A. and Serend, J.A. (1999). Training American listeners to perceive Mandarin tones. Journal of Acoustic Society of America 106 (6), 3649-3658.

[2] Halle, P.A., Chang, Y.-C. and Best, C.T. (2003). Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners. Journal of Phonetics 32, 395-421.

[3] Shih, C. and Lu, H.-Y.D. (2010). Prosody transfer and suppression: Stages of tone acquisition. Proc. Speech prosody, Chicago, IL.

[4] Hao, Y.-C. (2012). Second language acquisition of Mandarin Chinese tones by tonal and non-tonal language speakers. Journal of Phonetics 40, 269-279.

[5] Lin, T. and Wang, L. (2009). A Course on Phonetics. Beijing: Beijing University Press.

[6] Shen, H.H. (2006). Introduction to Standard Chinese Pinyin System. Beijing: Beijing Language and Culture University Press.

[7] Zhang, F. (2009). A study of pronunciation problems of English learners in China. Asian Social Science 5, 141-143.

[8] Jiang, L. and Allard, J. (2012). American Chinese learners’ acquisition of L2 Chinese affricates /ts/ and /ts(h)/. Journal of Acoustic Society of America 132 (3), 1936.

[9] Matsumoto, Y. (2011). Nihongo Bogo Washa ni Taisuru Chugokugo Hatsuon Kyoiku no Riron to Jissen [Chinese Pronunciation Training for Japanese Learners of Chinese: Theory and Practice], PhD thesis, Waseda University.

[10] Chiu, C.-Y., Kulls, D., Mixdorff, H., and Chen, S.-L. (2009). Preliminary perceptual evaluation of productions by German learners of Mandarin. Proc. NCMMSC, Lanzhou, China.