Articulatory and Acoustic Features of Mandarin /ɹ/: A Preliminary Study

Shuwen Chen[1], *Peggy Pik Ki Mok*[1]

[1]The Chinese University of Hong Kong
chenshuwen@link.cuhk.edu.hk, peggymok@cuhk.edu.hk

Abstract

Rhotic sounds in the world's languages have a wide range of variants, and are famous for their complexity in production. The current study examined the articulatory and acoustic features of Mandarin /ɹ/ using ultrasound imaging. The results showed that similar to English rhotics, Mandarin /ɹ/ could be articulated with various tongue shapes that were usually categorized as the bunched gesture (tongue tip pointing down) or retroflex gesture (tongue tip curling up). The variation between bunched and retroflex /ɹ/, however, was only found in the postvocalic and syllabic /ɹ/. Mandarin prevocalic /ɹ/ was produced with the tongue tip pointing down (bunched gesture). Acoustically, Mandarin /ɹ/ had a higher F3 than English /ɹ/ in the prevocalic and syllabic positions, and a higher F2 in the prevocalic position, indicating less rhoticity in Mandarin /ɹ/ than in English /ɹ/. Moreover, frication noise was often observed in the prevocalic /ɹ/, but not in all prevocalic tokens. Large interspeaker variation was found in using frication noise in the production of prevocalic /ɹ/.

Index Terms: rhotic sounds, Mandarin /ɹ/, acoustics, articulation, ultrasound imaging

1. Introduction

This is a phonetic study examining the acoustic and articulatory features of Mandarin /ɹ/. It has been well-studied that English /ɹ/ can be articulated with various tongue shapes with slightest acoustic consequence [1, 2, 3]. The continuum of tongue variation was usually roughly categorized as the retroflex and bunched gestures based on the tongue tip positions [2, 3]. It is unclear, however, if the articulatory variability is a language-specific phenomenon unique to English, or a universal phenomenon that can be found in other similar rhotic sounds, such as Mandarin /ɹ/. In addition, there has been a debate if Mandarin prevocalic rhotic sound is an approximant or a fricative in Mandarin phonology. The current study will provide phonetic data that could potentially answer these two questions.

Phonologically, Mandarin /ɹ/ can occur in the prevocalic, syllabic and postsyllabic positions. In the underlying forms, Mandarin /ɹ/ only exists in the prevocalic and syllabic positions. The /ɹ/ sound occur in the coda position after a r-suffixation process ('er-hua' 儿化). Together with the postalveolar fricative and affricates /ʂ/, /tʂ/ and /tʂh/, Mandarin prevocalic /ɹ/ are usually called the "retroflex consonants" in the literature and in the classroom settings [4, 5]. This is because early documentation of those consonants described those sounds as being articulated with the tongue tip curling up [4]. In the syllabic position, Mandarin /ɹ/ cannot be followed or preceded by any consonants. Therefore, there are only a few words that contains a syllabic rhotic in Mandarin Chinese, such as /ɹ̩/35 'son, child' ("儿"), /ɹ̩/214 'ear' ("耳"), and /ɹ̩/53 'two' ("二"). The r-suffixation process is a feature of Mandarin spoken in Northern China, such as Beijing, Shandong Province, and Hebei Province [6]. The r-suffix is a diminutive suffix, or used to refer to a familiar object [7, 8, 9]. Orthographically, it is represented by the word /ɹ̩ /35 'son, child' ("儿"). In Mandarin Chinese, one character represents one syllable in most cases. But for the character that combines with the word /er/35 after r-suffixation, the two characters will be pronounced as only one syllable. The syllabic /ɹ̩/ undergoes syllable contraction, and merges with the preceding vowels as part of the rime [10].

Previous studies demonstrated that there were acoustic and articulatory similarities between Mandarin and English /ɹ/. Acoustically, Mandarin /ɹ/ was characterized by a low F3 [12]. The frequencies of the first three formant of Mandarin /ɹ/ were around 372 Hz, 1459 Hz and 2118 Hz [11]. Articulatorily, Chao [4] stated that Mandarin rhotic sounds involved the tongue tip curling up. However, a recent study using palatography and EMA reported no retroflexion in the production of those consonants, and the constriction were reported to be around the post-alveolar region [12]. Lee [12] recorded 3 Beijing speakers (1 male and 2 females) with three EMA sensors on the speakers' tongue tip, anterodorsum, and posterodorsum. She found consistent F3 falling in acoustics, but no retroflex in the articulation was observed. Using ultrasound imaging, Gick, Campbell, Oh and Tamburri-Watt [13] examined the articulatory gestures of Mandarin prevocalic and postvocalic /ɹ/ produced by 1 native speaker. They found that Mandarin rhotics was produced with two active movements of the tongue – tongue anterior raising and tongue root backing. This was similar to the characteristics of English /ɹ/ articulation. They also observed that the tongue root backing gesture was around 33ms earlier in achieving maximum displacement than tongue anterior raising gesture. Given the similarities between Mandarin and English /ɹ/, it is interesting to examine if the articulatory variation can also be found in Mandarin /ɹ/.

One debate about the Mandarin prevocalic rhotic sound is whether this sound is an approximant or a fricative. This debate relates to the perceivable frication noise in some prevocalic /ɹ/ tokens. In the earliest description of Mandarin Chinese sound inventory, Karlgren [14] described Mandarin prevocalic rhotic sound as a fricative. Duanmu [5] also proposed that it should be categorized as a voiced fricative /ʐ/. He proposed that the relationship between Mandarin /ʂ/ and /ʐ/ is similar to that of English /s/ and /z/. Acoustic evidence of frication in prevocalic /ɹ/ has been reported in Smith [11]. This categorization, however, was criticized because it induced the only voiced fricative in Mandarin sound inventory. Other studies, following the tradition of Chao [4], described it as a post-alveolar approximant, and used /ɹ/ or /ɹ̺/ (an approximant that is produced with the apical part of the tongue) to represent the sound [10, 15, 16]. A detailed phonetic examination of Mandarin /ɹ/ is needed to shed light on the approximant-fricative debate. If frication noise was consistently found in prevocalic /ɹ/, it would provide strong evidence for the fricative account of prevocalic /ɹ/. If frication noise was an optional phonetic cue for prevocalic /ɹ/, the approximant account will be supported.

In summary, the current study aims to provide a systematic description of the acoustic and articulatory features of Mandarin /ɹ/ using ultrasound imaging. More specifically, the current study tried to answer two research questions: 1) if articulatory variation can be found in Mandarin /ɹ/; 2) if frication noise is a consistent phonetic cue for prevocalic /ɹ/.

2. Method

Twelve native Mandarin speakers (one male speaker and eleven female speakers) who speak a rhotic accent of Mandarin ("er-hua" in Mandarin) were recorded with ultrasound imaging. They were all born and grew up in Northern China (Beijing, Hebei and Shandong Province), so they naturally speak with a rhotic accent of Mandarin. Their average age was 23.33 years old (Range: 21-28, SD = 1.97). To compare Mandarin /ɹ/ with English /ɹ/, six native American English speakers (two male speakers and four female speakers) were recorded reading English words containing /ɹ/. The native English speakers had an average age of 20.83 (Range: 19-23, SD=1.34), and spoke a rhotic accent of English. The participants were asked to sit comfortably and speak with an ultrasound probe under their chins. The ultrasound data was collected with the EchoB ultrasound system together with the Articulate Assistant Advanced (AAA) software. Each speaker read the target words in a random order for 8 repetitions.

Mandarin stimuli included words containing prevocalic /ɹ/ coarticulated with the /ʅ a ɤ u/ vowels, postvocalic /ɹ/ with the /i ɿ ʅ y u a ɤ/ vowels, and syllabic /ɹ̩/ (see Table 1 for examples). The Mandarin words were produced with the carrier phrase / tʂɤ kɤ ___ ba/ "This is ___." (/ba/ is a sentence final particle in Mandarin). The English stimuli included words containing prevocalic, postvocalic and intervocalic English /ɹ/ coarticulated with the /ɑ æ ɛ ɪ ɔ u ʌ/ vowels, and syllabic /ɹ/. The /ɹ/ sound was embedded in different syllable positions – 47 prevocalic /ɹ/ in /#_V(C)/, /C_V(C)/ and /CC_V(C)/ words, 15 postvocalic /ɹ/ in /V_#/ words, 10 syllabic /ɹ/ in /C_C/, /C_#/ and /#_C/ words, and 3 intervocalic /ɹ/ in /C_C/, /#_C/, /C_#/ words. The target words were produced with the carrier sentence "what a ___ again" when the word started with a consonant, and "Speak of ___ again" when the word started with a vowel.

Table 1: *Sample words for the Mandarin stimuli.*

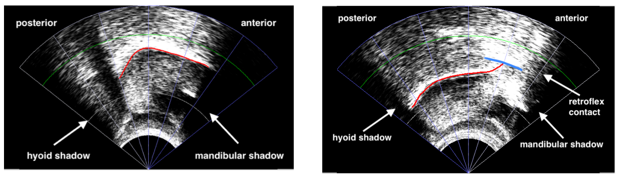
Syllable positions	Vowel contexts	Sample words
Prevocalic	/ʅ a ɤ u/	/ɻʅ/ 'sun' 日 /ɻɤ/ 'hot' 热
Postvocalic	/i ɿ ʅ y u a ɤ/	/sɹ/ 'thread' 丝儿 /yɹ/ 'fish' 鱼儿
Syllabic		/ɹ̩/ 'son' 儿 /ɹ̩/ 'ear' 耳

3. Results

3.1. Articulatory features of Mandarin /ɹ/

The ultrasound data showed that similar to English /ɹ/, Mandarin /ɹ/ can be produced with various tongue shapes. In order to simplify the discussion and enable cross-linguistic comparison with English /ɹ/, the tongue shapes were categorized as bunched and retroflex gestures. The representative raw ultrasound images of the bunched and retroflex tongue shapes were shown in Figure 1.

Two basic criteria were used in the categorization of bunched and retroflex gestures: 1) which part of the tongue was used to make the constriction, 2) whether the tongue tip is curling up or pointing down. While it was sometimes difficult to tell the position of the tongue tip based on a single ultrasound frame, the sequence of tongue contour movements from the preceding segment before /ɹ/ to following segment after /ɹ/ were examined as well. The author and another trained phonetician who have experience with ultrasound imaging did the categorization. They first did the categorization separately, and then discussed the different judgments together. If they had the same categorization, or they agreed with each other after discussion, the judgement of that particular token was marked as "same". If they disagreed with each other even after discussion, the judgement was marked as "different". The inter-rater reliability for all tokens was 93.98%. The results of the "same" categorization of the tongue shapes were summarized in Table 2, and the "different" tokens were discarded.



bunched tongue shape　　retroflex tongue shape

Figure 1: *Typical bunched and retroflex tongue shapes in Mandarin /ɹ/.*

Among the 12 speakers, 8 of them used the bunched gestures in all syllable positions, and 4 of them used both the bunched and retroflex gestures. There are mainly four patterns in the articulation of Mandarin /ɹ/. First, for all speakers, only the bunched gesture was used in the prevocalic position. The alternation of bunched and retroflex gestures was found only in syllabic and postvocalic positions. Second, each speaker used the same gesture (bunched or retroflex gesture) in the syllabic and postvocalic position. That is, if a speaker chose to use the bunched or retroflex gesture, he/she used that gesture in all vowel contexts in syllabic and postvocalic position. Third, the bunched gesture was more prevalent in the syllabic and postvocalic position (4 speakers used retroflex gesture and 8 speakers used bunched gesture). Fourth, the tongue shape of Mandarin /ɹ/ was not influenced by vowel contexts. In Mandarin, four vowels (/ʅ a ɤ u/) can be coarticulated with the prevocalic /ɹ/ and seven vowels (/i ɿ ʅ y u a ɤ/) can be coarticulated with postvocalic /ɹ/. The tongue shapes of Mandarin /ɹ/ were consistent in the same syllable position across the vowel contexts.

Table 2: *Summary of the tongue gestures by all speakers.*

Participants	**Prevocalic /ɿ a ɤ u/**	**Syllabic**	**Postvocalic /i ɿ ʅ y u a ɤ/**
P1	Bunched	***Retroflex***	***Retroflex***
P2	Bunched	***Retroflex***	***Retroflex***
P3	Bunched	Bunched	Bunched
P4	Bunched	Bunched	Bunched
P5	Bunched	Bunched	Bunched
P6	Bunched	***Retroflex***	***Retroflex***
P7	Bunched	Bunched	Bunched
P8	Bunched	Bunched	Bunched
P9	Bunched	Bunched	Bunched
P10	Bunched	***Retroflex***	***Retroflex***
P11	Bunched	Bunched	Bunched
P12	Bunched	Bunched	Bunched

3.2. Acoustic features of Mandarin /ɹ/

Frication noise was observed in many tokens of prevocalic /ɹ/, but never observed in syllabic and postvocalic /ɹ/. The data showed that the frication noise can be found in all vowel contexts in the prevocalic position. Figure 2 shows representative spectrograms of prevocalic /ɹ/ in different vowel contexts where frication noise could be found.

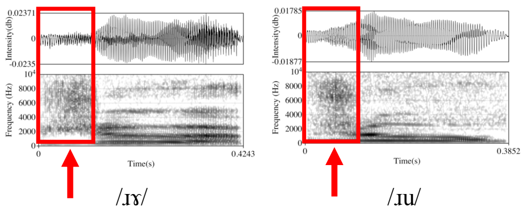


Figure 2: *Waveforms and spectrograms of prevocalic /ɹ/ in /ɤ/ and /u/ contexts by Speaker W10.*

Table 3 summaries observed frication noise in the /ʅ u ɤ a ɑ / vowel contexts for each speaker. Clearly, there was large interspeaker variation in using frication noise in the production of prevocalic /ɹ/. Some speakers produced frication noise in most prevocalic /ɹ/. For example, Speakers P6 produced frication noise consistently across all vowel contexts and in all repetitions. Speaker P7 and P11 produced frication noise in all tokens except some tokens in the /a/ and /u/ contexts. On the contrary, some speakers produced only a few tokens with the frication noise. For example, Speaker P3 consistently produced frication noise only when the prevocalic /ɹ/ was adjacent to the apical vowel /ʅ/. As can be seen from Table 3, frication noise was more often observed when /ɹ/ is followed by high vowels (/ʅ u/) than the two allophones [a ɑ] of the low vowel /a/. The presence of frication noise was influenced by tongue height rather than by categorical tongue shape differences – the retroflex or bunched tongue shape.

Table 3: *Summary of frication noise observed in Mandarin prevocalic /ɹ/ (+ indicates the presence of frication noise; – indicates absence of frication noise; "some tokens" means that frication noise could be found in some repetitions).*

Speakers	**/ʅ/**	**/u/**	**/ɤ/**	**/a/ [a]**	**[ɑ]**
P1	+	+	+	–	–
P2	some tokens	some tokens	+	–	some tokens
P3	+	–	some tokens	–	–
P4	+	+	+	some tokens	some tokens
P5	+	+	+	+	some tokens
P6	+	+	+	+	+
P7	+	+	+	+	some tokens
P8	–	some tokens	+	some tokens	–
P9	+	+	+	–	some tokens
P10	+	some tokens	some tokens	–	–
P11	+	some tokens	+	+	+
P12	+	+	+	–	some tokens

To examine the effects of Language (Mandarin /ɹ/ vs. English /ɹ/) and Syllable position (prevocalic, postvocalic and syllabic positions), the F3 and F2 of Mandarin and English /ɹ/ was transformed into Bark scale and compared. Linear mixed-effected models were performed on the F3 and F2 at the lowest F3 point of the /ɹ/ sound. The F3 and F2 of Mandarin and English /ɹ/ in prevocalic, postvocalic and syllabic positions were shown in Figure 3 and Figure 4.

Results showed that Mandarin /ɹ/ has a higher F3 and F2 than English /ɹ/ in the prevocalic and syllabic positions. The best model for F3 included Language, Syllable position, and the interaction between Language and Syllable position as fixed effects, and Participant and Items as random intercepts. The model results suggested that there was a main effect of Language and a main effect of Syllable position. The model also yielded a significant two-way interaction between Language and Syllable position. To understand the nature of the interaction, subsequent linear mixed-effects models were performed on formant values in each syllable position. In the prevocalic position, F3 of Mandarin /ɹ/ was significantly higher than that of English (Estimate = 1.510, Std. Error = 0.209, $t = -7.212$, $p < 0.001$). In the syllabic position, F3 of Mandarin /ɹ/ was also significantly higher than that of English (Estimate = 0.661, Std. Error = 0.214, $t = -3.083$, $p = 0.006$). No significant differences were found in postvocalic position for F3.

The best model for F2 included Language, Syllable position, and the interaction between Language and Syllable position as fixed effects, and Participant and Items as random intercepts. The model results showed that there was a main effect of Language, a main effect of Syllable position, and a significant interaction between Language and Syllable position. Subsequent analysis comparing English and Mandarin /ɹ/ in each syllable position showed that in the prevocalic position, the F2 of Mandarin /ɹ/ was higher than that of English (Estimate = 1.250, Std. Error = 0.278, t = -4.490, p < 0.001). No significant difference between the two languages in postvocalic and syllabic positions was found.

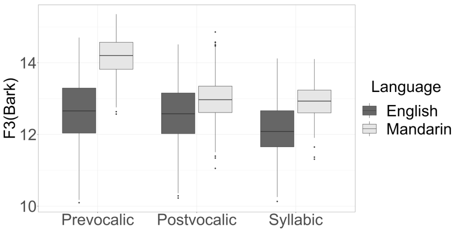


Figure 3: *F3 values of Mandarin and English /ɹ/ by native speakers.*

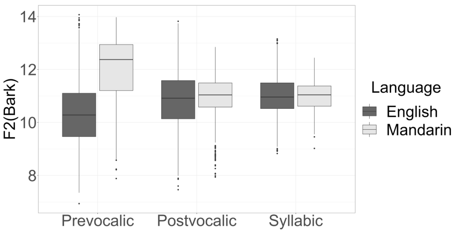


Figure 4: *F2 values of Mandarin and English /ɹ/ by native speakers.*

4. Discussion

Similar to English /ɹ/, Mandarin /ɹ/ is signaled acoustically by a low F3. However, the F3 of Mandarin /ɹ/ is higher than that of English /ɹ/ in prevocalic and syllabic positions, indicating that Mandarin /ɹ/ is less rhotic than English /ɹ/ in those two positions. In terms of articulation, the ultrasound data showed that Mandarin /ɹ/ could be articulated with multiple articulatory gestures. The articulatory variants found in Mandarin /ɹ/ were quite similar to articulatory gestures of English /ɹ/, and could be categorized as the bunched or retroflex tongue shape. The articulatory variability of rhotic sounds, therefore, was not a language-specific characteristic. It might be a universal property for rhotic sounds to have multiple realizations in articulation.

Secondly, segmental contexts were not found to be a factor that influence the choice of articulatory gestures in Mandarin. Previous studies suggested that in English, the retroflex gestures were preferred when the /ɹ/ sound was adjacent to low vowels or back vowels compared to high vowels and front vowels due to the compatibility of gestures [2]. The retroflexion involve retraction of the tongue body, which is also required for English low back vowel production. The lingual gestures of the bunched gesture involve the raising of tongue front, which is similar in English high front vowels. Therefore, the retroflex gesture is more compatible with low back vowels while the bunched gesture is more compatible with high front vowels in English. In Mandarin, however, articulatory gestures were not affected by vowel contexts. If a speaker chose to use a bunched or retroflex gesture, he/she would stick to that gesture despite the segmental contexts. The results suggested that lingual compatibility did not necessarily affect articulatory gestures in the production of the /ɹ/ sounds. The articulatory variation of Mandarin postvocalic and syllabic /ɹ/ was more related to individual preference in articulatory gestures.

Regarding the debate on Mandarin prevocalic rhotic sound, the data in the current study favors the account that Mandarin prevocalic rhotics should be categorized as an approximant. First, the frication noise was often observed in the prevocalic /ɹ/, but it did not occur in all prevocalic tokens. There was large interspeaker variation in using frication noise in the production of prevocalic /ɹ/. It suggests that frication noise is not an obligatory phonetic cue in the production of Mandarin prevocalic /ɹ/, at least for Mandarin spoken in Northern China. In addition, we found lowering of F3 in Mandarin prevocalic rhotic sound, which was the indicator of rhoticity of the /ɹ/ sound. It suggests that Mandarin prevocalic /ɹ/ shares acoustic characteristics of Mandarin postvocalic and syllabic /ɹ/ which was categorized as an approximant. Based on the data in the current study, it is more appropriate to categorize Mandarin prevocalic rhotic sound as an approximant rather than a fricative because it lacks the consistent presence of frication noise which is the most important phonetic cue for fricatives.

One limitation of the current study is that the number of male and female speakers is not balanced. The data of the current study did not show how gender might influence the tongue shape and individual variance. Future studies can be done to investigate the effect of gender on the production of Mandarin /ɹ/.

5. Conclusion

The current study examined the articulatory and acoustic characteristics of Mandarin /ɹ/. The ultrasound data showed that Mandarin /ɹ/ could be articulated with multiple articulatory gestures that were categorized as the bunched and retroflex gestures. Different from English /ɹ/ in which articulatory variation could be found in all syllable positions, Mandarin prevocalic /ɹ/ could be articulated with only the bunched gesture, while syllabic and postvocalic /ɹ/ could be produced with either the retroflex or bunched gesture. Also, the tongue shape of Mandarin /ɹ/ was not influenced by vowel contexts. Acoustically, Mandarin prevocalic /ɹ/ often involves fricative noises, but the fricative noise was not an obligatory phonetic cue in Mandarin prevocalic /ɹ/ production. Also, Mandarin /ɹ/ had a higher F3 value than English /ɹ/ in prevocalic and syllabic positions, indicating that the Mandarin /ɹ/ is less rhotic than English /ɹ/ in prevocalic and syllabic positions.

6. References

[1] P. Delattre, and D. C. Freeman, "A dialect study of American r's by x-ray motion picture," *Linguistics* 6, no. 44, pp. 29-68, 1968.

[2] J. Mielke, A. Baker, and D. Archangeli, "Variability and homogeneity in American English /r/ allophony and /s/ retraction," *Laboratory Phonology*, 10, pp. 699–730, 2010.

[3] X. Zhou, C.Y. Espy-Wilson, S. Boyce, M. Tiede, C. Holland, and A. Choe, "A magnetic resonance imaging-based articulatory and acoustic study of "retroflex" and "bunched" American English /r/," *The Journal of the Acoustical Society of America*, 123(6), pp. 4466–4481, 2008

[4] Y. R. Chao, *A Grammar of Spoken Chinese*. University of California Press, 1968.

[5] S. Duanmu, *The Phonology of Standard Chinese*. Oxford University Press, 2007.

[6] J. Wang, "The general rules of er-hua," Yuyan Wenzi Yingyong 儿化规范综论. 语言文字应用, *3*, 46–54, 2005.

[7] Y. Li, “The rhotic nucleus and rhotic phoneme in Mandarin Chinese,” *Yuwen Yanjiu.* 论普通话儿化韵及儿化音位. 语文研究, *59*, 21–26. 1996.
[8] B. Lin, “The rhotics in Mandarin Chinese,” *Yuyan Wenzi Yingyong*. 普通话的儿化. 语言文字应用, *4*, 91–94. 1992.
[9] J. Wang, “The norm of er-hua,” Yuyan Wenhua Yingyong. 儿化规范综论. 语言文字应用, *3*, 46–54, 2005.
[10] Y. H. Lin, *The Sounds of Chinese with Audio CD* (1 edition). Cambridge University Press, 2007.
[11] J. G. Smith, “Acoustic properties of English /l/ and /r/ produced by Mandarin Chinese speakers,” Master thesis. University of Toronto, 2010.
[12] W. S. Lee, “A phonetic study of the “er-hua” rimes in Beijing Mandarin,” *INTERSPEECH*, 2005
[13] B. Gick, F. Campbell, S. Oh, and L. Tamburri-Watt, “Toward universals in the gestural organization of syllables: A cross-linguistic study of liquids,” *Journal of Phonetics*, *34*(1), pp. 49–72, 2006.
[14] B. Karlgren, *Etudes sur la phonologie chinoise*, Vol. 5. E.-J. Brill, 1915-1926.
[15] M. Fu, “Phonemes and Pinyin symbols in the Beijing Speech,” Zhongguo Yuwen. 北京话的音位和拼音字母. 中国语文, *5*, 1956.
[16] W. S. Lee, and E. Zee, “Standard Chinese (Beijing),” *Journal of the International Phonetic Association*, *33*(1), 2003.