

A CORPUS STUDY OF RETROFLEX REALIZATIONS IN BEIJING AND TAIWAN MANDARIN

Yung-hsiang Shawn Chang

Department of Linguistics, University of Illinois at Urbana-Champaign, USA
yhchang@illinois.edu

ABSTRACT

This study investigated the spectral distributions of Mandarin retroflexion produced by Beijing Mandarin and Taiwan Mandarin speakers. Results showed that Taiwan Mandarin speakers' retroflex production had a higher spectral mean than their Beijing Mandarin counterparts, yet the two groups' distributions overlapped to a certain degree. This indicates that retroflexion is gradient and that retroflexion in Beijing Mandarin is acoustically more retroflexed. The degree of retroflexion was also modeled as a function of frication duration, as we hypothesized longer segmental duration allows fuller retroflex articulation.

Keywords: retroflexion, prosodic conditioning, Beijing Mandarin, Taiwan Mandarin, duration

1. INTRODUCTION

Lack of or a lesser degree of alveolar-retroflex contrasts has been claimed to be characteristic of Taiwan Mandarin and other southern dialects of Chinese (e.g., [5, 12]). However, recent studies have shown that retroflex neutralization in Taiwan Mandarin may not be a yes-no matter. Researchers reported realizations of Mandarin retroflexion to be influenced by vowel contexts [13], prosodic prominence [3], as well as sociolinguistic factors such as gender [3, 9] and the formality of tasks [9]. Assuming interim representations between the alveolars and retroflexes, a small number of studies initiated investigation in the gradience of retroflexion in Mandarin. For example, Li [13] described gradient alveolar-retroflex neutralization with spectral moments measures. Jeng [10] reported that the categorical perception of slightly retroflexed stimuli, as opposed to fully retroflexed and alveolar ones, incurred the highest response variance and the longest reaction time. In exploring the alveolar-retroflex acoustic continuum, this study asked whether retroflexion in Beijing Mandarin and Taiwan Mandarin occupied different places on this continuum. We looked at spontaneous speech, which more closely

reflected the speakers' day-to-day speech production, to examine the phonetic realizations of retroflexion by Beijing Mandarin and Taiwan Mandarin speakers.

Investigation on gradience of retroflexion is operationally difficult. In [9], different degrees of retroflexion were elicited by manipulating the formality of tasks—tokens extracted from wordlist readings were assumed to exhibit a greater retroflexion contrast than those from short answers to prompted questions. In [10], the subjects were instructed to artificially produce three levels of retroflexion. Given that segmental production is greatly conditioned by prosody (e.g., [2, 4]), Chuang & Fon [3] ToBI-labeled prosodic strength for their spontaneous data in terms of tonal realization, duration and amplitude in studying the effect of prosodic prominence on realizations of the alveolar-retroflex contrast. They categorized these acoustic properties based on some graded criteria. The current study continued investigations of gradience of retroflexion. Specifically, we examined whether duration, which is continuous data, affected the spectral characteristics of retroflexion. The acoustic measure of retroflexion was the first spectral moment (or COG, center of gravity), which is the most commonly used parameter to characterize fricatives of different places of articulation (e.g., [7]).

Kochanski & Shih's [11] Stem-ML model used weighted constraints to model phonetic realization, which was analyzed as a compromise between accuracy (prototypical phonetic form) and fluency (reduction for global optimization). This approach motivated a hypothesis in this study that Mandarin speakers may compromise retroflex articulation when the segmental duration was too short. That is, when a retroflex consonant is longer in duration, there will be more time for retroflexes to be fully articulated, and hence a lower COG. This hypothesis is articulatorily reasonable yet contradicts findings of several durational studies of alveolar-retroflex contrasts. For example, in Hindi, retroflex stop were found to have lower mean VOT

than alveolar counterparts [15]. Swedish dental /l/ appeared to be about twice longer than the retroflex /l/ [14]. Mandarin alveolars sibilants were reported to have longer frication than retroflexes do [9, 16]. If our hypothesis holds that longer duration allows more retroflex articulation, we have to explain why retroflexes are found to be longer than alveolar counterparts. The question of whether duration is a more robust parameter than spectral distribution for place discrimination or serves as a secondary perceptual cue will also arise. These concerns could be addressed by looking closely at the previous studies that reported shorter frication in Mandarin retroflexes. In [9] frication duration was averaged from that of fricatives and aspirated/unaspirated affricates. When discussed separately as in [6], frication from /ʃ/ was actually longer in duration than /s/ in most vowel contexts, whereas the frication from /tʃ, tʃʰ/ was shorter than that of /ts, tsʰ/ in all vowel contexts. Given the discrepancy between frication duration of fricatives and affricates, we predicted that either 1) frication duration is independent of realizations of Mandarin retroflexion, or 2) should frication duration be a perceptual cue in alveolar-retroflex distinction, it displays different directionalities for the perception of fricatives and affricates.

In summary, the goals of this study were to investigate the spectral distribution of retroflexion in Beijing and Taiwan Mandarin and whether realizations of Mandarin retroflexes were conditioned by frication duration.

2. METHODOLOGY

2.1. Tokens

The corpus used in this study consists of 185 hours of recordings from intermediate-advance level Chinese language classes at the author's affiliated institution. The spontaneous speech of eight Mandarin speaking teachers was analyzed. The speakers consisted of 4 female Beijing Mandarin speakers and 4 (3 female, 1 male) Taiwan Mandarin speakers. All of the Beijing Mandarin speakers lived in Beijing before coming to the United States; the Taiwan Mandarin speakers were all from northern Taiwan, speaking some or little Southern Min.

To investigate the natural gradience of retroflexion, we used the retroflex syllable that occurred most frequently in the corpus, *shi* /ʃl/ 'to be'. Using the same token across the database not only allowed us to hold the following vowel

context constant but also delimited the syntactic functions of the data, as segmental acoustics has been found to be heavily influenced by word class [1]. Thirty-five tokens of *shi* were randomly selected from each subject's recordings. Only 14 instances of *shi* could be found for one Taiwan Mandarin speaker; we included her data for analysis to balance the number of speakers between the Beijing and Taiwan dialects. All instances of *shi* that occurred at major prosodic junctures (i.e., the beginning/end of a sentence) were removed because they may exhibit different temporal characteristics. More than 80% of the tokens in this study were the second member of disyllabic compounds, such as *jiushi* 'that is' and *danshi* 'but'. In case of syllable contraction in forms of fricative deletion or voicing, they were excluded in the 35 tokens. Tokens occurring with noise in the background that would affect acoustic analyses were also excluded.

2.2. Analysis

The frication part of all *shi* tokens were labeled in Praat. The initial spectral analysis window was centered 10ms after the onset of frication. The subsequent windows were advanced in 10-ms intervals throughout the fricative. The spectral mean was derived from all tokens by executing a spectral moments analysis script in Praat.

Since our hypothesis on realizations of retroflexion was articulatorily motivated, we also took into consideration the effect of the surrounding segments. Only analyzing the *shi* data, we factored out any coarticulatory effect from the following segment. In previous modeling studies on Mandarin segments, a coarticulatory effect from the preceding segment was found to be insignificant (e.g., [16]). However, studies on retroflexion in other languages indicated that the preceding vowel has a larger coarticulatory effect on retroflexion than the following vowel—for which Hamann [8] claimed to be unique to retroflexion. Since we were unsure whether retroflex articulations in Beijing Mandarin and Taiwan Mandarin were comparable to each other or to other retroflex languages, we labeled the segment preceding /ʃ/ and included it as another factor on the realizations of *shi*.

3. RESULTS

The descriptive statistical results were reported for each individual speaker in Tables 1 and 2. Table 1

shows the mean duration of frication, the standard deviation (SD), and range. Table 2 presents the mean center of gravity, the SD and the range of spectral mean for all speakers. To further visualize the relationship between the temporal measure and spectral distribution of the fricative /ʃ/ as in *shi*, we collapsed all speakers' data into a scatter plot (see Figure 1).

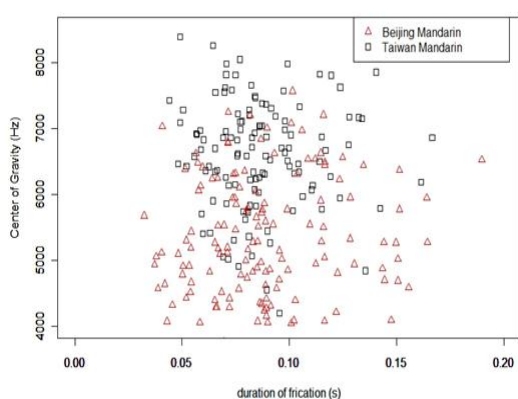
Table 1: Distributions of the frication duration (ms) for all speakers.

Beijing Mandarin	GJ	GYTA	JY	WYJ
range	32-150	40-156	50-164	41-190
mean (SD)	73 (25)	92 (31)	92 (29)	99 (34)
Taiwan Mandarin	WJH	ZYX	SJL	STJ
range	4.9-167	52-140	44-132	58-142
mean (SD)	94 (29)	82 (18)	86 (22)	86 (24)

Table 2: The center of gravity (kHz) of spectral distributions for all speakers.

Beijing Mandarin	GJ	GYTA	JY	WYJ
range	4.1-5.9	3.8-6.1	4.4-6.8	5.2-7.6
mean (SD)	4.7 (0.5)	4.8 (0.5)	5.7 (0.6)	6.4 (0.6)
Taiwan Mandarin	WJH	ZYX	SJL	STJ
range	6.1-8.4	4.2-8.3	5.3-7.5	5.0-7.0
mean (SD)	7.2 (0.6)	6.6 (0.9)	6.6 (0.6)	6.0 (0.5)

Figure 1: Relationship between the duration and the spectral mean of frication for two dialectal groups.



Putting Table 1 and Figure 1 together, we see that the mean durations of Beijing Mandarin and Taiwan Mandarin speakers' fricative production were not noticeably different. The similar temporal distributions allowed us to better compare the group COG distributions. As can be seen, Figure 1

shows partial distribution overlap between Taiwan Mandarin speakers' data and Beijing Mandarin speakers, with the former spreading to higher frequencies. Taiwan Mandarin speakers' data clustered between 6000-7500 Hz, which conformed to the range of COG of retroflex production reported in [3, 9, 10]. On the other hand, Beijing Mandarin speakers' data was centered around 5000 Hz. A Welch two-sample t-test showed a significant difference between the two groups ($t=-11.64$; $p<0.001$). Upon a quick summary, the phonetic realizations of retroflexion were not limited to a small frequency band. Since the data analyzed were all productions of *shi*, retroflexion in Beijing Mandarin and Taiwan Mandarin could be characterized as two clustered distributions that partially overlap on a continuum of retroflexion. In general, our Beijing Mandarin speakers' production had a lower COG, and hence exhibited a stronger degree of retroflexion than their Taiwan Mandarin counterparts.

In terms of phonetic context, we would especially like to see whether peripheral vowels or consonants with gestures that conflict with retroflex articulation (i.e., those that hold an anterior tongue position) have an effect on retroflexion. However, only /ə/ and /o/ were most frequently occurring segments before *shi* and the only segments that occurred more than once across all speakers' data. Since these two vowels were not articulatorily spaced enough, no consistent patterns of spectral distributions were revealed across speakers. Instead of plotting the COG and its corresponding preceding segment here, we will address the context effects after post-hoc analysis.

Correlation analyses were carried out on frication duration and COG of each individual's data. The results show that two Taiwan Mandarin speakers' (ZYX and WZY) data were weakly correlated (Pearson's $r = -0.26$ & -0.22 respectively). However, none of these correlation coefficients has reached a statistically significant level. To know if the preceding segment had an effect on the realizations of retroflexion, we further conducted a 2-way repeated-measures ANOVA, with two within-subject variables (frication duration and the preceding segment) and COG as the dependent variable, on each speaker's data individually. The results revealed no significant effect of preceding segments on COG for all eight speakers. However, frication duration did have a significant effect on COG for speaker ZYX

($F=5.41$; $p=0.03$). That is, for ZYX, COG was lower as the frication became longer in duration.

4. CONCLUSION AND DISCUSSION

This study investigated the spectral distributions of Mandarin retroflexion produced by Beijing Mandarin and Taiwan Mandarin speakers and whether frication duration could be a predictor of degrees of retroflexion. The results showed that Taiwan Mandarin speakers' retroflex productions had a significantly higher spectral mean than their Beijing Mandarin counterparts, although their distributions overlapped to a certain degree. This suggests that retroflexion is indeed gradient and that retroflexion in Beijing Mandarin may be acoustically more retroflexed than in Taiwan Mandarin. The parameter of duration did not well predict the degree of retroflexion for all speakers except for one Taiwan Mandarin speakers. We also looked at the effect of the preceding segment on the retroflex production and the effect was not statistically significant. The finding is in line with previous durational studies on Mandarin segments. However, the results should be interpreted with care as we did not have enough number of peripheral vowels and anterior consonants across all speakers' data to explore their coarticulatory effects on retroflexion.

We acknowledge that all our speakers in the corpus were language teachers and their speech might not be representative of the whole spectrum of Beijing Mandarin and Taiwan Mandarin speakers. Especially for some Taiwan Mandarin speakers who indeed make no alveolar-retroflex distinction, duration-conditioned retroflexion would not apply to their speech. While in this study a retroflexion continuum was constituted in terms of COG of frication, it will be interesting to investigate how the acoustic continuum ties into the perception of retroflexion. A following categorical perception experiment will allow us to probe into where the boundary is drawn between listeners of different dialects of Mandarin.

5. ACKNOWLEDGEMENTS

This study is made possible with the Critical Research Initiatives Grant (UIUC) and the NSF-funded project "DHB: An Interdisciplinary Study of the Dynamics of Second Language Fluency."

6. REFERENCES

- [1] van Bergem, D.R. 1993. Acoustic vowel reduction as a function of sentence accent, word stress, and word class. *Speech Communication* 12, 1-23.
- [2] Cho, T., McQueen, J.M. 2005. Prosodic influences on consonant production in Dutch: Effects of prosodic boundaries, phrasal accent and lexical stress. *Journal of Phonetics* 33, 121-157.
- [3] Chuang, Y.-Y., Fon, J. 2010. The Effect of Prosodic Prominence on the Realizations of Voiceless Dental and Retroflex Sibilants in Taiwan Mandarin Spontaneous Speech. *Proceedings of Speech Prosody 2010*.
- [4] Cole, J., Kim, H., Choi, H., Hasegawa-Johnson, M. 2007. Prosodic effects on acoustic cues to stop voicing and place of articulation: Evidence from radio news speech. *Journal of Phonetics* 35, 180-209.
- [5] Duanmu, S. 2000. *The Phonology of Standard Chinese* (1st ed.). Cambridge University Press.
- [6] Fong, L. 1985. Beijinghua yuliu zhong sheng yun diao de shichang (Duration of consonants, vowels, and tones in Beijing Mandarin speech). *Beijinghua Yuyin Shiyuanlu (Acoustics Experiments in Beijing Mandarin)*, 131-195.
- [7] Haley, K. L., Seelinger, E., Mandulak, K., Zajac, D.J. 2010. Evaluating the spectral distinction between sibilant fricatives through a speaker-centered approach. *Journal of Phonetics* 38, 548-554.
- [8] Hamann, S. 2003. *The Phonetics and Phonology of Retroflexes*. Ph.D. Dissertation, Utrecht: LOT Press.
- [9] Jeng, J.-Y. 2006. The acoustic spectral characteristics of retroflexed fricatives and affricates in Taiwan Mandarin. *Journal of Humanistic Studies* 40(1), 27-48.
- [10] Jeng, J.-Y. 2009. The Auditory Discrimination of Mandarin Retroflex Contrasts and Spectral Moment Analysis. *Chinese Journal of Psychology* 51(2), 157-174.
- [11] Kochanski, G., Shih, C. 2003. Quantitative measurement of prosodic strength in Mandarin. *Speech Communication* 41(4), 625-645.
- [12] Kubler, C. 1985. *The Development of Mandarin in Taiwan: A Case Study of Language Contact*. Taipei: Student Publishing Co.
- [13] Li, Y.-S. 2009. *Effects of Lexical Frequency and Neighborhood Density on Incomplete Neutralization in Taiwan Mandarin*. Ph.D. dissertation, CCU.
- [14] Livijn, P. 2002. Distribution of dental and retroflex l-sounds across some Swedish dialects. *Proceedings of 15th Swedish Phonetics Conference (FONETIK 2002)*.
- [15] Ohala, M., Ohala, J.J. 1992. Phonetic universals and Hindi segment duration. *International Conference on Spoken Language Processing*, Edmonton.
- [16] Shih, C., Ao, B. 1997. Duration study for the Bell Laboratories Mandarin Text-to-Speech System. In van Santen et al. (eds.), *Progress in Speech Synthesis*. New York: Springer.