Using Map Tasks to Investigate the Effect of Contrastive Focus on the Mandarin Alveolar-Retroflex Contrast

Yung-hsiang Shawn Chang, Chilin Shih

Department of Linguistics, University of Illinois at Urbana-Champaign, U.S.A.

yhchang@illinois.edu, cls@illinois.edu

Abstract

A series of map tasks were designed to investigate whether vowel contexts and prosodic prominence may affect the realizations of the alveolar and retroflex sibilants in Beijing Mandarin and Taiwan Mandarin. Results show that the realizations of the sibilants and their contrast varied by vowel context for most speakers, and were selectively subject to prosodic conditioning for only some speakers. The /u/ vowel context was where the alveolar-retroflex contrast was the least distinct and where an enhanced place contrast was achieved for half of the Beijing Mandarin speakers under focus condition. On the other hand, /u/ was the context where some Taiwan Mandarin speakers showed confusion over the place distinction or hypercorrected in response to focal prominence.

Index Terms: focus, vowel context, Mandarin alveolar and retroflex sibilants, Beijing Mandarin, Taiwan Mandarin

1. Introduction

There are five places of articulation for Mandarin fricatives and affricates. Among them, only the alveolar /s, ts, tsʰ/ and retroflex /ʂ, tʂ, tʂʰ/ sibilants have been reported to be subject to place neutralization in some southern dialects of Mandarin like Taiwan Mandarin [9]. In Taiwan Mandarin (hereafter TM), a number of contextual factors such as vowel contexts [11] and prosodic prominence [4] have been identified to greatly influence the magnitude of the alveolar-retroflex contrast. Extralinguistic factors like gender [4] and formality [8] were also found to be sources of phonetic variability in realizing the alveolar and retroflex sibilants in TM. In contrast to the variability reported for TM, a distinct place contrast for Beijing Mandarin (hereafter BM), regardless of contextual factors, is generally suggested in the literature on Mandarin phonology (e.g., [6]). However, in view of observations made in recent sociolinguistic studies (e.g., [3]) that the alveolar-retroflex neutralization is in progress in BM, particularly among the younger generation of speakers, this study aims to investigate whether the same phonetic variability is also present in sibilant realizations in BM. In examining the Mandarin alveolar-retroflex contrast in various phonetic contexts, this study will provide a detailed acoustic characterization of variation in Mandarin sibilant realizations by dialect, speaker and vowel context.

1.1. Spectral Properties of Mandarin Alveolar and Retroflex Sibilants

How the spectral energy is distributed in the frication noise can effectively distinguish sibilants of different places of articulation. Wu & Lin [14] argued that the primary acoustic cue distinguishing Mandarin alveolars from retroflexes resides in the position of the lowest spectral prominence, with /s/ higher than /ʂ/. Stevens et al. [13] further pointed out that the spectrum of frication noise in Mandarin /s/ is centered at high frequencies (around the frequency of F5), compared to a prominent peak in the F3 region for /ʂ/ and some weaker noise energy in the F2 region. Besides characterizing local differences in the fricative spectrum, places of articulation can also be differentiated by the measure of center of gravity (COG). COG is calculated as the mean of frequencies weighted by intensity when the power spectrum is considered to be a probability distribution [10].

1.2. Factors Affecting the Realizations of Mandarin Alveolar/Retroflex Sibilants and Their Contrasts

The acoustic realization of a phonological contrast can be influenced by many factors, among which the effects of prosodic prominence on strengthening phonological distinctiveness have been well studied in the literature. For consonants, Cole et al. [5] reported that the voicing contrast is generally enhanced for English stops in accented condition, although they also found diminishing voicing contrast under accent. While expecting increased acoustic distinctiveness under accent, Cole et al. [5] actually found the place of articulation contrast between English /p, t, k/ to be uniformly strengthened instead of being enhanced under accent. Their results suggest that contrast enhancement may not be the only result that prosodic prominence gives rise to. Similar variability has also been reported in Chuang & Fon [4] for Mandarin sibilant production in accented vs. unaccented conditions. Specifically, two different strengthening strategies in the accented condition were observed in their subjects: 1) enlarge the place contrast by producing the alveolar or retroflex sibilants or both toward a more extreme direction (i.e., retroflexes having a lower COG; alveolars having a higher COG), and 2) strengthen the place distinction by increasing the center of gravity of both sibilants.

Realizations of alveolar [12] and retroflex [7] sibilants are both subject to the vowel effect. For Mandarin sibilants, Jeng [8] reported a lower COG measured over the fricative followed by a rounded vowel than by a non-rounded vowel. In addition, the vowel context also seems to be an important factor in the magnitude of the Mandarin alveolar-retroflex contrast, as Li [11] found a smaller COG difference between the two sibilants before high vowels for TM speakers.

Another factor that may affect Mandarin alveolar and retroflex production is word frequency. High-frequency words are produced more quickly than low-frequency words. In addition, high-frequency forms tend to induce phonetic reduction [1]. Given that almost every Chinese syllable has a number of different lexical representations, one has to control syllable frequency as well as consider the homophone effect in lexical retrieval in production. Therefore, in this study, all the homophones of a target syllable (i.e., an alveolar or a retroflex syllable) were pooled in the calculation of lexical frequency.

In summary, while impressionistic accounts generally describe BM as exhibiting a consistent alveolar-retroflex distinction, it is unclear whether the same phonetic variability found in TM exists in BM as well. This study is in particular interested in the variability of Mandarin sibilant realizations in

response to vowel contexts, prosodic prominence and their interaction. To this end, three research questions are addressed: 1) Do the realizations of the two sibilant categories and the alveolar-retroflex contrast vary across vowel contexts in both BM and TM? 2) Do these phonetic realizations vary in different prosodic conditions for both BM and TM? 3) Is there any interaction effect between the vowel contexts and prosodic prominence?

2. Methods

2.1. Subjects

The subjects included 4 BM speakers (2 males, 2 females) and 6 TM speakers (4 males, 2 females). All of the BM speakers were born and raised in Beijing; the TM speakers were all from northern Taiwan. While 5 out of the 6 TM speakers speak some Taiwan Southern Min, Mandarin was reported to be the major language used in their families. All subjects were students at the University of Illinois, aged between 20 and 30 years, and had not resided in the United States over 5 years.

To exclude the population of TM speakers who make no alveolar-retroflex distinction, a screening test was conducted before data analysis. The screening materials came from speakers' reading the Mandarin version of "The North Wind and the Sun". All of the alveolar and retroflex tokens were later extracted for an identification task conducted on two native speakers (1 BM speaker and 1 TM speaker). In view of previous literature [9] that reported alveolars as the default form for Mandarin alveolar-retroflex contrast, only the accuracy of retroflex productions was computed. It was decided that a speaker's retroflex productions must be correctly identified by the two listeners at least 60% of the time in order to be judged capable of making the place distinction. Two male TM speakers' data were excluded by this screening test. This study will proceed with data from 4 BM speakers and 4 TM speakers.

2.2. Map Tasks and Stimuli

All audio recordings were conducted in a sound-attenuated booth in the Phonetics Lab at the University of Illinois. An AKG C520 head-worn condenser microphone was used to record acoustic signals of the subjects' speech onto a Marantz PMD570 recorder.

To obtain balanced (i.e., an equal number of alveolar vs. retroflex tokens in various vowel contexts) yet natural speech data for analysis, a series of 5 map tasks were used to elicit alveolar and retroflex productions. The instructions were written at the top of each map: 1) Your partner (i.e., the investigator) is lost and has a misoriented map. Based on your map (see Figure 1 for a partial map for the subjects), you need to correct your partner's inaccurate statements of geographical relations. 2) After the location of all the sites has been confirmed, you should decide on a route that stops at every site on the map and then describe the route to your partner. The subject and the investigator were seated across from each other with a plastic board situated in between such that the two people could not see each other's map.

There were 60 target stimuli (2 places of articulation * 2 focus conditions * 3 vowel contexts /a, i, u/ * 5 samples—of the same place-vowel combination varying by tone and manners of articulation) for the map tasks. Focused productions were target syllables uttered in a corrective manner (i.e., subjects corrected geographical terms on a map), and unfocused productions were those uttered in a non-corrective manner (i.e., subjects gave the route on a map).

All stimuli were 4-syllable pseudo-words. The target stimuli were composed of one alveolar/retroflex syllable, a non-alveolar/retroflex syllable and a disyllabic compound. The disyllabic compounds are meaningful—they were all landscape terms (e.g., river) or scenic spot names (e.g., park); the first two syllables were pseudo names of the landscape term or scenic spot name. Each constituent of the 4-syllable pseudo-words was checked against the SUBLEX-CH corpus [2] and shared similar log word frequencies. The fillers differed from the target stimuli only by the first syllable; the first syllable of fillers was not an alveolar or retroflex syllable.

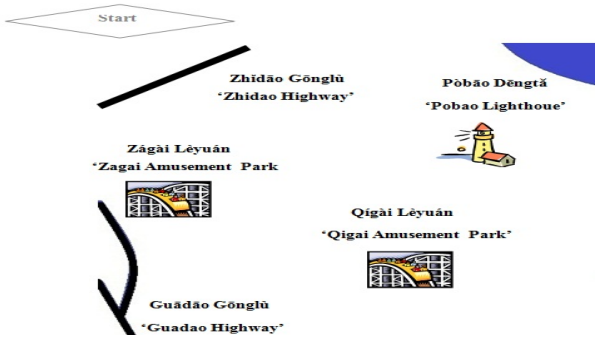

Figure 1: *A partial map for the subjects*

3. Results

3.1. Segmentation and Acoustic Measurement

The sound files along with their text transcription were sent to automatic segmentation using the PennPhonetics Lab Forced Aligner [15]. In order to measure the COG of frication, all the target syllables were further re-segmented in Praat to separate the frication from the stop closure and aspiration in cases of affricates. COG was then obtained from the middle 30 ms of the frication interval, to which a high-pass filter set at 1000 Hz was applied to eliminate low frequency noise.

Besides investigating how Mandarin alveolar and retroflex sibilants are realized across vowel contexts and between focus and non-focus conditions, this study is also interested in whether the phonological contrast is conditioned by these two factors. The contrast was measured in terms of the difference between the COG values of the two sibilant categories (i.e., COG of alveolars minus that of retroflexes). Since the alveolar sibilants exhibit higher COG values than their retroflex counterparts, a greater positive COG difference is expected in case of contrast enhancement.

3.2. Realizations of Contrastive Focus

This study used the duration measure of the target syllables to determine whether contrastive focus was implemented. Each subject's duration data was submitted to a paired T-test. The results show that all subjects' focused productions were significantly longer than their unfocused counterparts. That is, contrastive focus was manifested in duration.

3.3. Statistical Analysis

To get a general picture of what factors are most important in predicting the COG measure, the sibilant production data was

fitted with a linear mixed-effects model, with focus, vowel, places of articulation, gender and dialect as fixed factors, and the preceding segment and subject as random factors. The statistical analysis revealed a significant main effect for place of articulation ($p<.0001$), gender ($p<.05$), and vowel ($p<.001$). More specifically, retroflex sibilants, male productions, and productions in the /u/ context have a significantly lower COG.

To further investigate within-subject and between-subject variation in sibilant realizations, a multiple linear regression analysis was conducted on each subject's data. The results are summarized in Table 1. No main effect of focus was found in the analysis of any subject's data. The results, however, did reveal a significant effect of vowel for all speakers, except TM3. In case of a significant interaction effect, the data were divided into subsets to investigate the interaction effect. For BM speakers, the analysis for BF1 and BF2 shows that within each vowel subset, place was still found to be a significant factor, although it has the least magnitude of effect on sibilant productions in the /u/ context. In addition, a significant focus effect ($p<.05$) was found for BF2 in her alveolar productions in the /u/ context. For BM2, within each vowel subset, place was still found to be a significant factor. A significant effect of focus was also found for his alveolar/retroflex+/a/ syllables ($p<.05$) and retroflex+/u/ syllables ($p<.01$). As for TM speakers, the analysis for TF2 shows that she only made an alveolar-retroflex distinction in the /i/ context ($p<.0001$). For TM1, the results revealed a statistically significant difference between the alveolar and retroflex production in the /a/ context ($p<.01$) as well as a significant focus effect in retroflex productions in the /a/ context ($p<.05$). For TM3, a significant effect of focus was observed for retroflex productions in the /a, i/ contexts ($p<.05$), but not for alveolar ones.

dialect	gender	subject	place	vowel	interaction
BM	female	BF1	***	*	*** (place×vowel)
		BF2	***	*	*** (place×vowel)
	male	BM1	***	**	
		BM2	***	**	*(place×focus) *(place×vowel)
TM	female	TF1	***	***	
		TF2	***	***	*** (place×vowel)
	male	TM1		***	*(focus×vowel) *(place×vowel)
		TM3	***		* (place×focus)

Table 1: *The statistical results for all speakers (*indicates p<.05; ** p<.01; *** p<.001)*

Upon examining the effects of vowel and focus on the realizations of the alveolar-retroflex contrast, another linear mixed-effects analysis was conducted, with COG difference as the dependent variable. The results revealed a significant main effect for gender ($p<.05$), vowels ($p<.001$) and dialect ($p<.05$). To quantify the place contrast made across vowel contexts and focus conditions, a 3x2 repeated-measure ANOVA was conducted on each subject's data, with COG difference as the dependent variable, and vowel and focus as the independent variables. Figures 2-4 are plots of the mean of each subject's place contrast data by vowel, with a 95% confidence interval.

For BM speakers, the results of BF1 and BM1 show that their alveolar-retroflex contrast varied significantly by vowel context, but not by focus condition. For BF2's alveolar-retroflex contrast, in addition to the effect of vowel, a post-hoc test indicates that her place contrast was subject to the effect of focus in the /u/ context. As for BM2, his place contrast was subject to both the effects of vowel and focus. When compared to his sibilants realizations, it should be noted that a greater place contrast was variably achieved in different vowel contexts for BM2: the enlarged contrast in the /u/ context was made by enhancing the retroflexed sibilants, whereas the enlarged contrast in the /a/ context was made by producing the two sibilants towards a more extreme COG direction.

For TM speakers, the results of TF1 and TM1 show that their place contrasts did not vary significantly by vowel or focus. As for TF2, her place contrast varied significantly by vowel, but not by focus. TM3's alveolar-retroflex contrast was influenced by different vowel contexts and focus conditions. When compared to his realizations of alveolar and retroflex sibilants, it should be noted that a greater place contrast was variably achieved in different vowel contexts for TM3: In enlarging the place contrast, TM3 only enhanced his retroflex production in the /a, i/ contexts (realized as a lower COG) in response to contrastive focus, while the realizations of his alveolar production were not subject to the same focus effect.

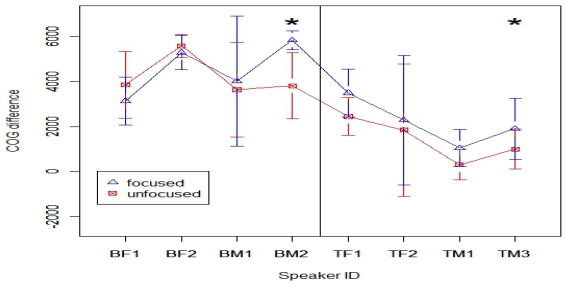

Figure 2: *Place contrast in the /a/ context*

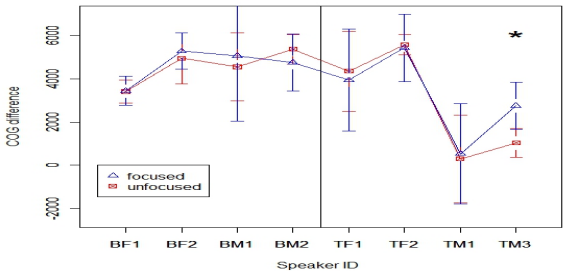

Figure 3: *Place contrast in the /i/ context*

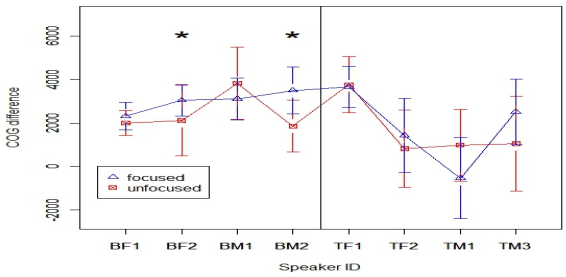

Figure 4: *Place contrast in the /u/ context*

4. Discussion and Conclusion

In this study, it was found that place, vowel, and gender best predict the COG of all subjects' sibilant realizations. The alveolar-retroflex contrast was found to vary by dialect in addition to place, vowel and gender. That said, the realizations of individual sibilant categories were not found to be significantly different by dialect. However, as far as the place contrast is concerned, the BM speakers in this study did make a significantly larger contrast than their TM counterparts.

The presence or absence of focus on target syllables did not affect the realizations of Mandarin alveolar and retroflex sibilants or the contrast between them when the eight speakers' data was grouped. While the duration data shows that focus was indeed implemented, the insignificance of the focus effect suggests that 1) focal prominence was not uniformly implemented within and across speakers, and 2) focal prominence was not manifested in the realizations for both or any of the sibilant categories. A close look at each individual's data revealed great within- and between-subject variability in sibilant realizations as well as various contrast strengthening strategies in response to focal prominence. Particularly, focus interacted with vowel in affecting sibilant productions in that half of the BM speakers signaled focal prominence in the /u/ context (and the /a/ vowel context for BM2). In contrast, two Taiwan Mandarin speakers (TM1 and TM3) showed an enhanced contrast in the /a/ vowel context and one (TM3) in the /i/ context. It should be noted that different strategies were also observed for the speakers' contrast enhancement. For example, while only retroflex+/a/ syllables were chosen to enhance for TM3 and alveolar+/a/ syllables for TM1, BM2 realized both alveolar and retroflex sibilants in the /a/ context towards a more extreme direction in enhancing the contrast. As for enhancement of the place contrast in the /u/ context, BF2 chose alveolar sibilants to enhance and BM2 chose retroflex sibilants.

Among all speakers, TM1 and TF2 stood out as having almost neutralized the alveolar and retroflex sibilants in most vowel contexts. Upon a close listen to their sound files, many of TM1's retroflex productions sounded like alveolar ones, and most of TF2's /retroflex+a/ and /retroflex+u/ productions seemed to be replaced by the alveolar counterparts regardless of focus condition. While both speakers had passed the screening test and were considered able to make the place distinction, they behaved quite differently in the map tasks. It is speculated that TM1 and TF2's alveolar-retroflex contrast may be partially lexicalized, meaning that they have encoded certain words with a retroflex, but in cases of uncertainty they default to alveolars. In addition, TM1 sometimes exhibited hypercorrection in response to focal prominence, as evidenced by negative COG difference values for his place contrast in the /u/ context.

In sum, the BM speakers in this study produced a significantly larger alveolar-retroflex contrast than the TM speakers, although the realizations of alveolar and retroflex sibilants respectively were not found to be significantly different by dialect. The realizations of alveolar/retroflex sibilants and the contrast varied by vowel context for most of the speakers and were variably subject to prosodic conditioning for some speakers. In response to focal prominence, the place contrast was enhanced in certain vowel contexts for some speakers. Specifically, half of the BM speakers enhanced the sibilants in the /u/ context where the place contrast tends to be acoustically weakened by coarticulation. On the other hand, the other two vowel contexts were where place enhancement was observed for TM speakers. Consistent with what Chuang & Fon [4] reported for TM, some BM and TM speakers were found in this study to realize either or both sibilant categories differently when a situation calls for a better place distinction. In addition, this study also found that enhancement may not occur independently of vowel contexts for BM as well as TM speakers. Taken all the results together, it is suggested that 1) realizations of Mandarin alveolar and retroflex sibilants are variable, 2) enhancement of the place contrast in response to focus can be variably achieved across different vowel contexts and speakers, and 3) the /u/ context is the environment where alveolar-retroflex neutralization is more likely to occur.

5. Acknowledgements

This study is made possible in part with a Cognitive Science/ Artificial Intelligence Award (Beckman Institute, UIUC).

6. References

[1] Bybee, J. (2002). Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language Variation and Change*, 14, 261-290.

[2] Cai, Q., & Brysbaert, M. (2010) SUBTLEX-CH: Chinese Word and Character Frequencies Based on Film Subtitles. *PLoS ONE* 5(6): e10729.

[3] Chen, P. (1999). *Modern Chinese: history and sociolinguistics*. Cambridge University Press.

[4] 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. In *Speech Prosody 2010 Proceedings*, Chicago.

[5] 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.

[6] Duanmu, S. (2000). *The phonology of Standard Chinese*. Cambridge University Press.

[7] Hamann, S. (2003). *The phonetics and phonology of retroflexes*. Utrecht: LOT.

[8] Jeng, J.-Y. (2006). The acoustic spectral characteristics of retroflexed fricatives and affricates in Taiwan Mandarin. *Journal of Humanistic Studies*, 40(1), 27-48.

[9] Kubler, C. C. (1985). The influence of Southern Min on the Mandarin of Taiwan. *Anthropological Linguistics, 27*(2), 156-176.

[10] Forrest, K., Weismer, G., Milenkovic, P., & Dougall, R.N. (1988). Statistical analysis of word-initial voiceless obstruents: Preliminary data. *JASA,* 84, 115–124.

[11] Li, Y.-S. (2009). *Effects of Lexical Frequency and Neighborhood Density on Incomplete Neutralization in Taiwan Mandarin*. Ph.D. dissertation. CCU.

[12] Niebuhr, O., Clayards, M., Meunier, C., & Lancia, L. (2011). On place assimilation in sibilant sequences—Comparing French and English, *Journal of Phonetics*, 39(3), 429-451.

[13] Stevens, K., Li, Z., Lee, C.-Y., & Keyser, S.J. (2004). A note on Mandarin fricatives and enhancement. *From Traditional Phonology to Modern Speech Processing*.

[14] Wu, Z., & Lin, M. (1989). *Overview of Experimental Phonetics*. Beijing: Higher Education Press.

[15] Yuan, J., & Liberman, M. (2008). Speaker identification on the SCOTUS corpus. In *Proceedings of Acoustics 2008*, 5687-5690.