The Effect of Prosodic Prominence on the Realizations of Voiceless Dental and Retroflex Sibilants in Taiwan Mandarin Spontaneous Speech

Yu-Ying Chuang, Janice Fon

Graduate Institute of Linguistics, National Taiwan university, Taiwan
d98142003@ntu.edu.tw, jfon@ntu.edu.tw

Abstract

This study investigated how voiceless dental and retroflex sibilants were realized in response to prosodic prominence in Taiwan Mandarin. It has been indicated that retroflex sibilants in Taiwan Mandarin were replaced with their dental counterparts due to the lack of retroflex phonemes in Min. The spontaneous speech of eight Mandarin-Min bilinguals was analyzed. Results showed that whether sibilants were merged or distinguished was highly dependent on gender and region, both of which are factors corresponding to the frequency of using Min. Moreover, the strengthening effect was found to be achieved by various strategies. In general, speakers who showed sibilant distinction enlarged or maintained the contrast in the prominent condition. As for those who merged the two sibilants, dental sibilants were advanced in place of articulation for signaling prominence.

Index Terms: prosodic prominence, dental sibilant, retroflex sibilant, Taiwan Mandarin

1. Introduction

Accented or stressed syllables always create the percept of prominence. In pursuit of its underlying mechanism, most of early studies focus on suprasegmental cues. Specifically, it has been indicated that pitch, duration, and amplitude are three well-established correlates of prominence, both at lexical and phrasal levels. Such a correspondence was observed in a number of languages. Take Mandarin for example. Chao [3] suggested that sentence stress is manifested by pitch range enlargement, duration lengthening, and increase of loudness. Nevertheless, in addition to suprasegmental cues, more and more recent studies shifted attention to phone realizations at segmental levels. For instance, Beckman et al. [2] reported that English /ɑ/ is articulated more rapidly and involves larger jaw movements when it occurs in the accented syllable. Incorporating both articulatory and acoustic data, Cho [4] further indicated that vowel features can be separately enhanced, depending on the source of prominence. In general, it has been shown that greater articulatory efforts are devoted in the prominent condition, and therefore vowels are usually more peripherally and distinctively realized.

As for consonants, previous studies also reported that in the accented position, the voicing contrast between obstruents is enlarged [5]. It should be noted, however, that contrast enhancement is not necessarily the only result that accent can give rise to. For instance, when studying stop realizations in English radio news speech, Cole et al. [6] examined the effect of accent on two phonological contrasts, voicing and place of articulation, by taking the same set of acoustic measurements (VOT, F_0, closure duration, burst amplitude). They found that whereas the voicing contrast is generally enhanced in the accented condition, the place of articulation contrast, nevertheless, is not enlarged, but uniformly strengthened. For

certain acoustic correlates, diminished contrasts are even observed in some of the speakers. Given their results, it seems plausible to derive that the effect of accent is far from homogenous. Instead, it should be dependent on a number of factors, such as types of phonological contrast, acoustic correlates, speaker variability, etc. To further pursue this issue, the present study intends to investigate the effect of prosodic strengthening, with respect to the contrast between voiceless dental and retroflex sibilants in Taiwan Mandarin.

There are three pairs of voiceless dental and retroflex sibilants in Mandarin, including unaspirated affricates /ts/ and /tʂ/, aspirated affricates /tsʰ/ and /tʂʰ/, and fricatives /s/ and /ʂ/. In particular, articulatory data showed that dental and retroflex sibilants in Mandarin differ crucially in place of articulation, with dental sibilants having a more anterior constriction position than retroflex sibilants [9]. Nonetheless, one aspect regarding dental and retroflex sibilants in Taiwan Mandarin should be noted; that is, the status of retroflex sibilants is debatable. The main reason is that about 80% of the population in Taiwan are bilingual speakers of Mandarin and Min, a southern variety of Chinese languages. In Min, there is the same set of dental sibilants as in Mandarin, but no retroflex counterparts. In consequence, Taiwan Mandarin is often characterized as having no retroflex sibilants, for they are frequently substituted by dental ones [7].

However, the lack of retroflex sibilants in Taiwan Mandarin is not widely accepted by researchers. Later studies showed evidence that dental and retroflex sibilants are distinguished in the acoustic domain [10]. A closer examination revealed that such a discrepancy might have resulted from speaker and context differences. Specifically, in previous studies different speaker groups were targeted at. It should be noted that although most Taiwan Mandarin speakers are Mandarin-Min bilinguals, their frequency of using Min is not necessarily similar. In fact, the use of Min in Taiwan is disproportional: usually males are more willing to speak Min, and speakers in the south of Taiwan speak Min much more often than those in the north [1]. If the frequency of using Min is influential, we should expect speakers differing in the frequency of using Min to have discrepant sibilant production. As for contextual difference, earlier impressionistic studies mostly derived their observations from natural speech, whereas later acoustic studies applied acoustic measurements to laboratory speech. Since speakers are usually aware of their pronunciations under experimental settings, laboratory speech might not be able to fully reflect their actual articulation. In this regard, in the present study we look into spontaneous speech, which should more authentically represent the production of sibilants in Taiwan Mandarin.

Aside from better controls regarding speaker and context, the examination of prosodic strengthening effect in this study also shed light on the merging issue of sibilants in Taiwan Mandarin as well. By comparing sibilant realizations in the unaccented and accented conditions, we could have the

following predictions. In the unaccented condition, dental and retroflex sibilants could be either distinguished or merged. If sibilants were distinguished, the distinction should be maintained or possibly exaggerated in the accented condition, consistent with the prosodic strengthening effect found in previous studies. On the contrary, if sibilants were merged when unaccented, there would be two possibilities in the accented condition. For one, dental and retroflex sibilants might show distinction when accented. This finding would therefore imply that the merge of dental and retroflex sibilants is only superficial. That is, speakers still have two separate sibilant phoneme categories, which are only contrasted in the strengthening condition. However, if dental and retroflex sibilants were not distinguished in the accented position either, then it implies that the merge is an underlying one, in which speakers only have dental phonemic representation for both sibilants in their production. In this regard, the effect of enhancement would not act upon sibilant contrast, but solely on dental sibilants.

In summary, there are two specific goals in this study. First, we would like to disentangle the merging issue of dental and retroflex sibilants in Taiwan Mandarin. Spontaneous speech is examined, and speakers are divided into different groups for controlling their frequency of using Min. Second, the effect of prosodic prominence on sibilant realizations is pursued. How the prosodic strengthening effect is achieved and its relation to sibilant merging in Taiwan Mandarin is particularly of interest.

2. Method

2.1. Participants

Eight (4 males, 4 females) fluent Mandarin-Min bilinguals were recruited. Half of the speakers were born and grew up in Taipei, the northern metropolitan city in Taiwan; the other half were from Kaohsiung, the southern metropolitan city. All of them were aged from 20 to 35 at the time of recording.

2.2. Data collection

Each speaker had a 30-minute interview with one interviewer in Mandarin. The interviewer would ask open questions related to speakers' personal experience or opinions towards certain issues, and speakers responded to all questions spontaneously. Frequently discussed topics included favorite food, movies, traveling experience, government policy, etc.

2.3. Analysis procedure

2.3.1. Sibilant labeling and measurement

All voiceless dental and retroflex sibilants in the recording were labeled and measured. The main frication part of each sibilant was identified; for affricates, the burst was excluded. Sibilant spectra from the 10 ms in the middle of frication, considered to be the most steady portion, were extracted. One modification was done for the aspirated affricate pair /ts^h/ and /tʂ^h/. Instead of the 10 ms in the middle of the frication, we took spectra from the first one-third portion of the frication, in order to minimize the effect of aspiration, usually carrying formant information of the following vowels. Finally, we filtered out the frequency range below 1500 Hz to eliminate low frequency noise, and calculated each sibilant's centroid frequency, which has been identified as an acoustic parameter

that reliably distinguishes the place of articulation for voiceless sibilants [10] [11].

2.3.2. Prominence labeling

We adopted a modified version of Pan-Mandarin ToBI [12] to label the prominence levels for each syllable. There are four levels in total, from S0 to S3. S3 is for the prominent condition, S0 and S1 are the reduced ones, and S2 is the default condition. Nonetheless, unlike the original criteria that solely rely on tonal realizations, we included two additional parameters for the identification of prominence levels, i.e., duration and amplitude. The modified set of criteria is provided in Table 1.

Table 1. *Modified criteria for prominence labeling.*

Stress	Tone	Amplitude	Duration
S0	lexically neutral tone	soft	shortened
S1	loss of original tonal shape	soft	shortened
S2	Default	default	default
S3	tone expanded/raised	loud	lengthened

2.3.3. Data pruning

Several controls were made before conducting statistical analyses. Because we are currently interested in the effect of prosodic prominence on sibilant realizations, only S2 (the default/unaccented condition) and S3 (the accented condition) tokens were analyzed. Furthermore, vowel context was controlled. It has been shown that rounded vowels would greatly lower the centroid frequency of preceding sibilants due to the coarticulation effect [10]. Therefore, only sibilants in the unrounded vowel context were included for analyses, mainly because only in the unrounded vowel context could we collect sufficient tokens for dental and retroflex sibilants at both accent levels. Finally, word class to which each sibilant belongs was identified. This was done because whether a word is a content word or function word would affect articulation [13]. To avoid such a confounding effect, in this study we merely focused on sibilant tokens that occurred in the content word category. Table 2 shows the distribution of S2 and S3 sibilant tokens that are followed by unrounded vowels and belong to the content word category for each speaker.

Table 2. *The distribution of sibilants for all speakers (TP: Taipei; KH: Kaohsiung).*

Region	Place	Gender							
		Female				Male			
		XHR		XHY		HSK		YYS	
		S2	S3	S2	S3	S2	S3	S2	S3
TP	Dental	88	39	83	34	105	17	123	22
	Retroflex	350	101	296	116	508	61	413	70
		CJH		CN		KCZX		LZW	
	Place	S2	S3	S2	S3	S2	S3	S2	S3
KH	Dental	49	13	63	10	38	13	73	30
	Retroflex	192	41	311	26	272	58	225	96

3. Results

Figure 1 presents all speakers' averaged spectra of dental and retroflex sibilants in the S2 and S3 conditions. The mean centroid frequency is shown in Figure 2. A *place* (2) × *stress*

(2) two-way ANOVA was conducted separately for each speaker. The results are summarized in Table 3.
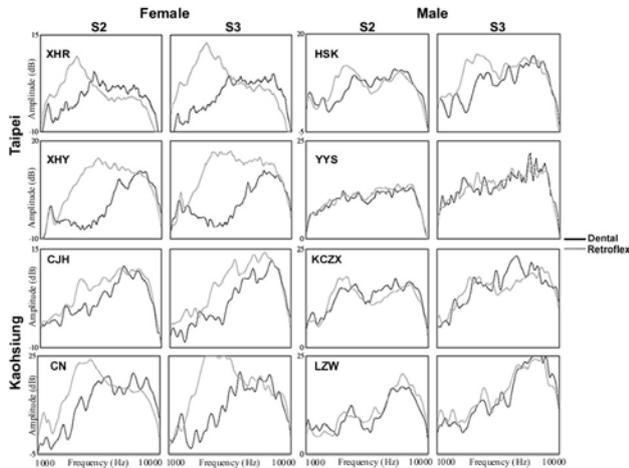

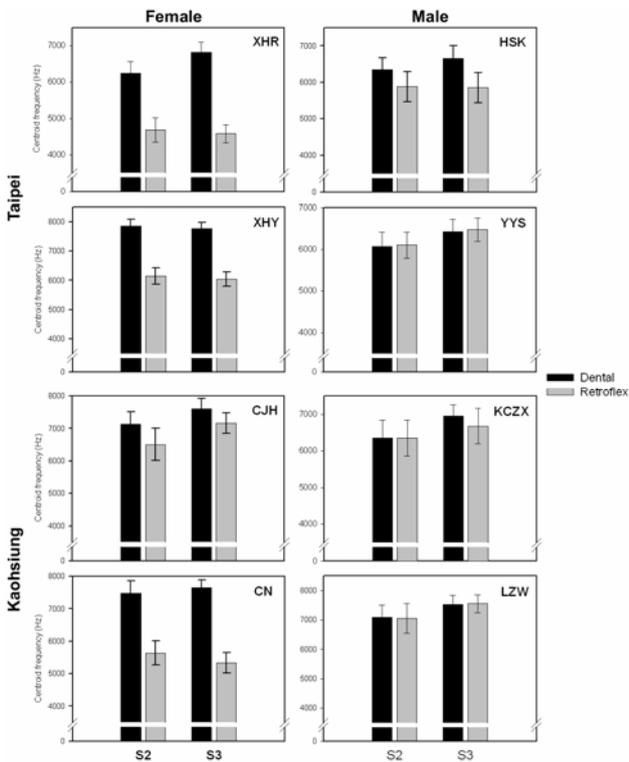Figure 1: *The averaged spectra of all subjects.*

Figure 2: *The mean centroid frequency of all subjects.*

Table 3. *The statistical results for all speakers.*

Region	Gender	Speaker	Stress	Place	Stress × Place
Taipei	Female	XHR	**	**	**
		XHY		**	
	Male	HSK		**	
		YYS	**		
Kaoshiung	Female	CJH	**	**	
		CN		**	
	Male	KCZX	*		
		LZW	**		

Note. * indicates significance. *: *p* < .01; **: *p* < .001.

As can be seen, the factor *place* was significant for five out of eight speakers. Specifically, in the S2 condition, all

female speakers realized dental sibilants with significantly higher centroid frequency than retroflex sibilants. On the other hand, three out of four male speakers did not significantly distinguish dental and retroflex sibilants.

The factor *stress* was reported significant for five speakers. In particular, how dental and retroflex sibilants were realized in the S3 condition was correlated with whether the two sets of sibilants were distinguished in the S2 condition. For speakers who showed clear sibilant distinction, there were three types of realizations for prominence. First of all, three speakers did not show significant *stress* effect, meaning that their dental and retroflex sibilants were not differently realized in the S2 and S3 conditions. This was particularly so for Speaker XHY. The second prominent realization was that sibilant contrast was enlarged in the S3 condition. For Speaker XHR, post hoc *t*-test revealed that her dental sibilants had significantly higher centroid frequency in the S3 condition (*p* < .001). It could also be observed from Figure 2 that Speakers CN and HSK also showed larger sibilant contrast in the S3 condition by realizing retroflex and dental sibilants toward more extreme directions respectively, though the difference was not significant. The last type of prominent realization was found for Speaker CJH. Particularly, both her dental and retroflex sibilants had higher centroid frequency in the S3 than in the S2 condition. The degree of sibilant contrast, nevertheless, was not enhanced.

As for speakers who did not significantly distinguish dental and retroflex sibilants in the S2 condition (Speakers YYS, KCZX, LZW), it could be seen that their stress realizations were consistent. That is, in the S3 condition, both sibilants were realized with significantly higher frequency. In addition, dental and retroflex sibilants were not significantly distinguished either in the S3 condition.

4. Discussion

The present study examined the effect of prosodic prominence on the realizations of dental and retroflex sibilants in Taiwan Mandarin. In particular, we investigated the realizations of both sibilants in the unaccented and accented conditions in spontaneous speech. Results showed that Taiwan Mandarin speakers differed in their production of sibilants, which in turn affected how the prosodic strengthening effect was achieved in the accented condition.

With respect to sibilant realizations, gender played a crucial role. In the unaccented (S2) condition, all female speakers showed a clear distinction of dental and retroflex sibilants; on the contrary, all male speakers except one failed to successfully distinguish the two sibilants. Moreover, the effect of region where speakers were from was effective as well, though playing a more minor role. Particularly for male speakers, neither of the two speakers from Kaoshiung, in southern Taiwan, distinguished retroflex sibilants from dental ones. Taken together, the findings in this study indicated that whether dental and retroflex sibilants were merged or not was highly correlated with the frequency of using Min. As noted above, males have higher frequency of using Min than females, and Min is more often spoken as one goes further south. In this regard, higher frequency of using Min introduces more profound influence from Min, and the merging of the two sibilants is therefore more likely to occur. In addition to the frequency of using Min, the robust gender effect could also be accounted for from the sociolinguistic perspective. It has been indicated that females are more conservative speakers than males; as a result, males usually use more nonstandard forms than females [8]. Such a hypothesis was also verified in the

current study; that is, males apparently had higher degrees of sibilant merging rate than females.

As for the effect of accent, our results were generally consistent with what Cole et al. [6] had found in their study. To be precise, the prosodic strengthening effect on dental and retroflex sibilants in Taiwan Mandarin was indeed variably achieved. In particular, speakers adopted different strategies to create prominence for their sibilant production. For those five speakers who normally showed sibilant contrasts in our study, three strengthening strategies were identified. The first strategy was that no accent difference was induced in the prominent condition. Such a result might imply that the strengthening effect was not realized on the enhancement of the acoustic correlate for the place of articulation contrast. It was likely that the strengthening effect would be shown on other articulatory parameters, such as the shape of the tongue.

The second strategy was to enlarge the contrast between dental and retroflex sibilants. The enhancement of phonological contrasts is the most typical strengthening effect as mentioned in previous research. The current study further found that the enlargement of sibilant contrast in Taiwan Mandarin was more likely to be unidirectional than bidirectional. Specifically, instead of realizing both dental and retroflex sibilants toward more extreme directions, Taiwan Mandarin speakers only enhanced either of the two sibilants to create larger contrasts. Moreover, dental sibilants were more often chosen to be enhanced than retroflex ones. The last strengthening strategy was the uniform strengthening effect for both dental and retroflex sibilants. That is, the spectral centroid for both sibilants was increased to the same extent, without creating larger contrast, as shown in the case of Speaker CJH. In line with the second strengthening strategy, it seemed that to signal prominence, most Taiwan Mandarin speakers would increase sibilants' spectral centroid frequency: some speakers did it only for dental sibilants (strategy 2), whereas others did it for both sibilants (strategy 3). Possibly it was because higher spectral centroid could give rise to more obvious high-frequency noise sounding effect, thus creating the perception of prominence.

With regard to the three male speakers who did not distinguish dental and retroflex sibilants, it was shown that both sibilants were consistently articulated with higher centroid frequency in the prominent condition. Just as in the unaccented condition, sibilants still appeared to be merged even when accented. This finding implicated that the merging phenomenon for these speakers was an underlying one, instead of a surface one. In other words, the substitution of dental sibilants for retroflex ones was the norm for their natural speech, and their strengthening strategy for dental sibilants was to increase spectral frequency, indicating a more anterior place of articulation. Since these speakers were proficient in Min and also used Min frequently, it was very likely that this sibilant merging was induced by Min. It was also possible that the enhancement of sibilants' spectral correlates found in Mandarin was borrowed from the strengthening strategy used in Min, considering the fact that Min only has dental sibilants. Further research on Min data will be necessary in order to confirm such a speculation.

5. Conclusion

This study examined dental and retroflex sibilants in Taiwan Mandarin and their realizations in the prosodically prominent condition. Specifically, results indicated that both gender and region were determinant factors for sibilant realizations. Such a finding inferred that the frequency of using Min was highly correlated with sibilant realizations in Taiwan Mandarin. Furthermore, it was found that the strengthening effect of prosodic prominence was not homogeneous across speakers in spontaneous speech. In the accented condition, speakers who showed sibilant distinction generally enlarged the contrast by producing either of the two sibilants towards a more extreme direction, or they strengthened both sibilants by increasing spectral frequency, without contrast enhancement. On the contrary, speakers who ordinarily had no sibilant contrast did not distinguish sibilants either in the accented condition. This suggested that their substitution of dental sibilants for retroflex ones was an underlying process. In consequence, the prosodic strengthening effect was achieved by enhancing the acoustic correlates of dental sibilants.

6. References

[1] Ang, U. (1997). 1997 Taiwan gonggong changsuo shiyong yuyan diaocha [A survey on language use in public occasions in Taiwan, 1997]. In C. Tung (Ed.), *Taiwan yuyan fazhan xueshu yantaohui lunwenji [Proceedings of the Conference on Language Development in Taiwan]* (pp. 83-100). Hsinchu: Quanmin Bookstore.

[2] Beckman, M. E., & Edwards, J. (1994). Articulatory evidence for differentiating stress categories. In P. A. Keating (Ed.), *Papers in Laboratory Phonology III: Phonological structure and phonetic form* (pp. 7-33).

[3] Chao, Y.-R. (1968). *A Grammar of Spoken Chinese.* Berkeley and Los Angeles: University of California Press.

[4] Cho, T. (2005). Prosodic strengthening and featural enhancement: Evidence from acoustic and articulatory realizations of /ɑ, i/ in English. *Journal of Acoustical Society of America, 117*(6), 3867-3878.

[5] 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.

[6] 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.

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

[8] Labov, W. (1990). The intersection of sex and social class in the course of linguistic change. *Language Variation and Change, 2*, 205-254.

[9] Ladefoged, P., & Wu, Z. (1984). Places of articulation: An investigation of Pekingese fricatives and affricates. *Journal of Phonetics, 12*, 267-278.

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

[11] Jongman, A., Wayland, R., & Wong, S. (2000). Acoustic characteristics of English fricatives. *Journal of Acoustical Society of America, 108*(3), 1252-1263.

[12] Peng, S., Chan, M. K. M., Tseng, C., Huang, T., Lee, O. J., & Beckman, M. E. (2007). Toward a Pan-Mandarin system for prosodic transcription. In S.-A. Jun (Ed.), *Prosodic Typology: The Phonology of Intonation and Phrasing.* Oxford: Oxford University Press.

[13] van Bergem, D. R. (1993). Acoustic vowel reduction as a function of sentence accent, word stress, and word class. *Speech Communication, 12*, 1-23.