GEOGRAPHICAL VARIATION OF THE MERGING BETWEEN DENTAL AND RETROFLEX SIBILANTS IN TAIWAN MANDARIN

Yu-Ying Chuang[1], Ching-Chu Sun[1], Janice Fon[2], and R. Harald Baayen[1]

Eberhard-Karls University of Tübingen[1], National Taiwan University[2]
yu-ying.chuang@uni-tuebingen.de ;ching-chu.sun@uni-tuebingen.de; jfon@ntu.edu.tw;
harald.baayen@uni-tuebingen.de

ABSTRACT

This study investigates the geographical distribution of pronunciation variation of voiceless dental and retroflex sibilants in Taiwan Mandarin. Previous studies indicated that the merging of the two sibilants is geographically dependent [17, 6]. However, the geographical effects in these studies are not easy to interpret due to the limited number of speakers and regions. For the current study, we recruited 331 native speakers of Taiwan Mandarin from 120 regions in Taiwan. In a picture-naming task, 30 dental/retroflex-initial words were elicited from each speaker. The data were analysed with Generalized Additive Mixed models [26]. The analysis revealed a robust effect of geographical location, with merging being less common in metropolitan cities as compared to the surrounding areas.

Keywords: Geographic variation, merging, dental sibilant, retroflex sibilant, Taiwan Mandarin

1. INTRODUCTION

When one language comes into contact with another, changes in one or both languages inevitably occur [21]. These contact-induced changes can be observed at various linguistic levels, and phonological or phonetic change is usually one of the phenomena that are most likely to happen [19]. The issue of how the sound systems of two contact languages interact with each other is sociolinguistically interesting, because social factors often play a crucial role in determining how sounds change [15]. This study examines the variation of sibilant merging in Taiwan Mandarin from a sociolinguistic perspective.

Taiwan is a highly bilingual society. In addition to the official language Mandarin, the majority of people in Taiwan speak an additional substrate language. As more than 70% of the people in Taiwan are ethnically Min, most people in Taiwan are Mandarin-Min bilinguals [9]. The extensive contact between the two languages has caused the sound system of Mandarin to be substantially influenced by Min [14, 11]. One of the well-documented phonetic variations in Taiwan Mandarin is the merging between dental and retroflex sibilants.

In Mandarin there are three pairs of voiceless dental and retroflex sibilants, including the fricatives /s/ and /ʂ/, the affricates /ts/ and /tʂ/, and the aspirated affricates /tsʰ/ and /tʂʰ/. Min, on the other hand, has the same three dental sibilants but none of the retroflex counterparts. The absence of retroflexes in Min is commonly considered as the cause of deretroflexion (i.e., producing less retroflexed sibilants or even no retroflexes) in the speech of Mandarin-Min bilinguals [14, 11, 7]. Deretroflexion results in neutralization of the two sibilants in Taiwan Mandarin. This merging phenomenon has been extensively studied in relation to various social factors. For example, males tend to have a higher degree of sibilant merging than females [23, 12], and merging is more frequently observed in spontaneous speech than read speech [18]. Other social predictors that show an effect on the degree of merging include age, educational level, and socio-economic status [17, 18].

One aspect of the merging variation that has not yet been fully explored is its geographical distribution. The lack of interest in the geographical effect on sibilant merging might be due to the fact that such effects are highly correlated with the influence of Min. Higher frequency or proficiency of Min usually leads to more severe merging [17, 18, 20]. Since Min frequency of use and Min proficiency is geographically determined, with more Min use and higher Min proficiency in the south than in the north [1, 10], it is therefore often assumed that sibilant merging follows the same geographical distribution as frequency or proficiency of Min. To the best of our knowledge, only two studies so far have directly addressed this issue [17, 6]. Although the findings of these studies are generally consistent with the assumed geographical effect, the numbers of participants and locations in both studies are however limited: Lin [17] recruited 45 participants and divided them into Taipei (the capital) and non-Taipei groups, whereas Chuang [6] examined 15 speakers from ei-

ther Taipei or Kaohsiung (the biggest city in the south). Given the small sample size in these studies, a more comprehensive and larger-scale investigation is necessary.

The goal of this study is to examine how the merging variation between dental and retroflex sibilants in Taiwan is geographically distributed. Instead of dividing speakers into separate location groups, we quantify geographical location by measuring the longitude and latitude of the region for each speaker. We then use Generalized Additive Mixed models (GAMM) [26] to model the non-linear interactions between them, following the approach of Wieling et al. [24] and Wieling et al. [25]. It is hypothesized that if the geographical effect coincides with the effect of Min influence, merging should aggravate from the north to the south. On the contrary, if the geographical effect is inconsistent or at odds with the effect of Min influence, a different geographical distribution of the merging variation should emerge.
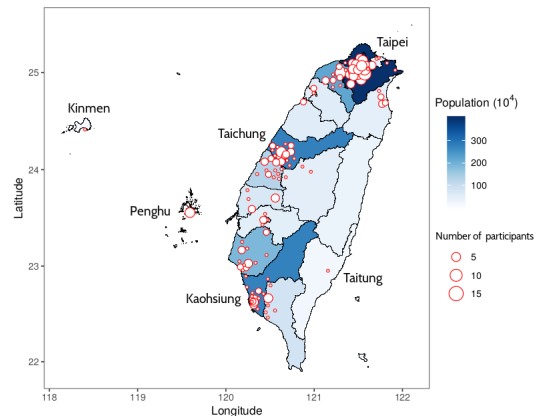
2. METHODS

2.1. Participants

A total of 331 native speakers of Taiwan Mandarin (194 males and 188 females) that are ethnically Min were recruited for the experiment. The age of the participants ranged from 18 to 30 (mean = 22.16). The geographical location of each participant was coded according to self-reported residential information, and was operationally defined as the place where the participant had lived for the longest period of time before the age of 18. In total 120 distinct locations were identified.

Figure 1 presents the geographical distribution of the participants in this study, along with the population size of each city coded in the background[1]. As can be seen, the population density is higher in the west of Taiwan. This is due to the fact that the east of Taiwan predominantly consists of mountainous areas. The major cities in the west of Taiwan are Taipei, Taichung, and Kaohsiung, respectively. Consistent with the population distribution, 67% of our participants were from these three major cities. Almost half of them (48%) were from Taipei, where the data were collected.

2.2. Materials and procedure

Thirty dental/retroflex-initial Mandarin words (fifteen each) were selected. Eight of the words start with fricatives (/s/ or /ʂ/), ten start with affricates (/ts/ or /tʂ/), and twelve start with aspirated affricates (/tsʰ/ or /tʂʰ/). Moreover, thirteen words have their

Figure 1: The geographical distribution of population in Taiwan, and the participants in this study.
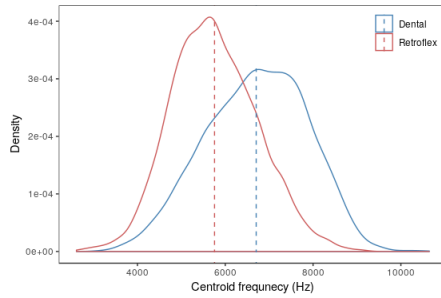
initial sibilants followed by rounded vowels. For the rest of the words, the sibilants are followed by unrounded vowels. Except for the initial position, no dental or retroflex sibilants occur in the rest of the words. Another set of thirty words were selected and included as fillers. These filler words do not contain any dental or retroflex sibilants.

We conducted a picture-naming task to elicit sibilant production. The pictures of all sixty words were presented in a sequential manner. Participants were asked to name the pictures naturally and spontaneously. Whenever the identification of a picture was not successful, the experimenter would provide oral hints to elicit the correct production. When the production of a word was unclear or accompanied by other vocal activities such as laughing or swallowing, repetitions were requested. All participants filled in a language background questionnaire about their residential history and language use. On a likert-scale from one to seven, they self-rated their fluency in Min and the frequency of using Min.

2.3. Acoustic measurement

To quantify sibilant realizations, we adopted the measurement of centroid frequency (henceforth CenFreq), a reliable measure for distinguishing different places of articulation for sibilants [13]. Specifically for studying the dental-retroflex distinction in Mandarin, CenFreq is widely used to measure their realizations as well [12, 16, 5].

For each sibilant, the frication part was manually labeled at zero-crossing boundaries. We then took a 30 ms spectral slice in the middle of each frication. Given that the realization of aspirated affricates includes aspiration, the middle part of an aspirated affricate could potentially coincide with aspiration. To prevent this, the 30 ms spectral slice of the two as-

Figure 2: The distribution of centroid frequency (Hz) of dental and retroflex sibilants in this study. The dashed lines indicate the respective means.

pirated affricates was taken from the first half of the sibilant realization. We filtered out frequencies below 1000 Hz before computing CenFreq. All labeling and measuring was done in Praat [4].
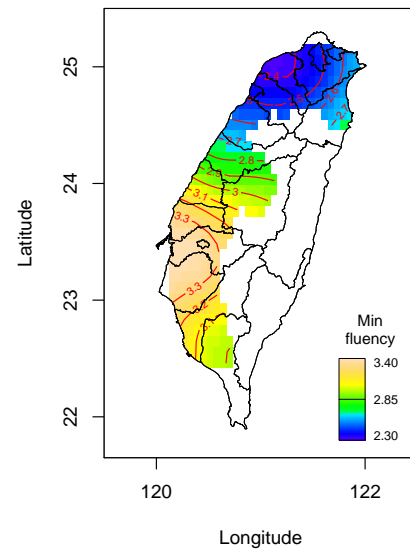
CenFreq is negatively correlated with the length of the front cavity prior to the constriction point. A more dental production results in higher values of CenFreq, whereas a more retroflex production gives rise to lower values of CenFreq. The distribution of CenFreq for dental and retroflex sibilants is presented in Figure 2, which plots the CenFreq distribution of dental and retroflex sibilants in this study.

3. RESULTS

In total 11,364 sibilant tokens were collected, with repetitions included. The data of eight participants were excluded, because the corresponding geographical regions were outliers in the current data set (six from Penghu, one from Kinmen, and one from Taitung, see Figure 1). Tokens that were produced with too short frication, with wrong pronunciation, and with accompanying vocal activities were excluded prior to analysis (about 4%). This left us with a total of 10,616 sibilant tokens, 5,262 of which are dentals and 5,354 are retroflexes.

3.1. Min proficiency and frequency of use

We first examined whether our participants' Min proficiency and frequency of use follows the same geographical distribution as described in previous studies [1, 10]. As rating is a measure of ordinal scale, we thus fitted an ordinal GAMM [3] to Min proficiency ratings, with gender and the non-linear interaction between longitude and latitude as predictors. Results showed that while the effect of gender did not reach significance ($p = 0.68$), there was however a tendency suggesting the geographical effect on Min proficiency ($p = 0.058$). As can be seen in Figure 3, Min proficiency varies along the latitudi-

Figure 3: Geographical variation of Min fluency.

nal scale, which gradually increases from the north to the south. This pattern of results confirms that speakers in the south speak Min more fluently than speakers in the north [10].
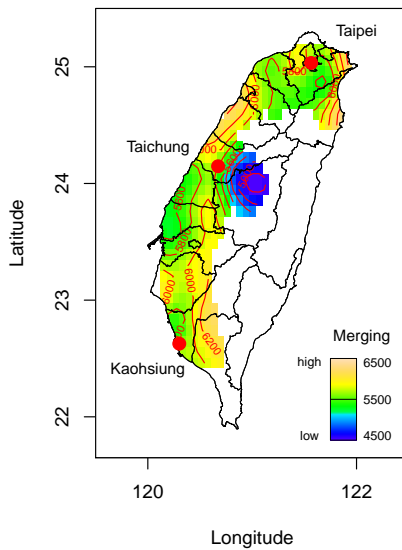
For Min frequency of use, we fitted another ordinal GAMM with the same predictors. However, neither the effect of gender, nor the effect of geographical location reached significance in this model. We suspect that the absence of the geographical effect is due to the fact that we collected the data in one location (Taipei). This may have led to a similar frequency of Min use among our speakers.

3.2. Sibilant merging

To examine geographical effects on sibilant merging, we again used a GAMM. The residuals of a canonical GAMM, however, violated the normality assumption. We therefore used a quantile GAMM [8] instead to relax such a restriction. We modeled the degree of sibilant merging by substracting the non-linear interaction between longitude and latitude for dentals from the corresponding interaction for retroflexes. We added gender and vowel context (rounded/unrounded) as fixed-effect control variables. Finally, we included by-participant and by-word random intercepts into the model.

The model revealed significant effects for both gender and vowel context (gender: $p < .001$; vowel context: $p < .0001$). As expected, CenFreqs were lower for males (due to longer vocal tracts). In addition, the following rounded vowels also led to lower CenFreqs (as lip protrusion lengthens the vocal tract). The interaction between gender and vowel

Figure 4: Geographical variation of sibilant merging. The three red points indicate the city centers of Taipei, Taichung, and Kaohsiung.

context was significant as well ($p < .001$), suggesting an asymmetrical lowering effect of the rounded vowel context for male and female speakers.

Crucially, geographical location influenced sibilant merging to a considerable extent ($p < .00001$), the effect of which is presented in Figure 4. With the geographical effect of dentals being controlled for, high CenFreqs (less retroflexed pronunciation) of retroflexes indicate a higher degree of merging (coded by warm colors), while low CenFreqs indicate a lower degree of merging (coded by cold colors). In contrast to the geographical pattern of Min fluency (see Figure 3), merging does not increase latitudinally. Instead, the degree of merging seems to be determined more by the degree of urbanization.

In Figure 4, we mark the urban areas of the three major cities (Taipei, Taichung, Kaohsiung). As can be seen, the degree of merging is relatively consistent in the three urban areas. Outside these areas, the merging situation is more variable. Generally speaking, the degree of merging increases when one moves away from the three metropolitan cities. The only exception to this pattern of results is a decrease of merging on the eastern side of Taichung (the blue area in the center of Figure 4). A closer examination of the data revealed that to the east of Taichung, there are only two female speakers, both of whom happen to be very retroflexed speakers. One of them is actually the most retroflexed speaker among all our participants. It is thus suspected that the decreased merging in this area may primarily be driven by speaker idiosyncrasy.

4. GENERAL DISCUSSION

This study investigated the geographical variation of the merging between dental and retroflex sibilants in Taiwan Mandarin. Different from previous studies where only limited number of speakers and locations were examined [17, 6], a larger group of speakers from more evenly distributed geographical locations was recruited for this study. In addition, by quantifying geography with longitude and latitude, we examined sibilant merging in a two-dimensional space. Such a method provided us with a clearer picture of how variation is distributed and propagated both vertically and horizontally.

One major finding of this study is that merging variation is, to a large extent, independent of the influence of Min: pronunciations of speakers with a higher proficiency in Min do not necessarily undergo more extensive merging. Instead, the results reported here suggest that the degree of merging is strongly tied to urbanization. Generally speaking, merging is relatively uncommon in urban areas, and the degree of merging increases in the surrounding suburban or rural regions. This result is reminiscent of Trudgill's gravity model [22], which states that the diffusion of a given linguistic change is dependent on population size. Thus change is usually spread to big cities first and from there it gradually diffuses to smaller ones. Although whether or not sibilant merging originates in the big cities in Taiwan remains an empirical question, a diffusion pattern is clearly observed. The gravity model, however, does not explain why more merging is found outside the cities, instead of inside the cities, as it would have predicted that cities are the center of variation and change.

One potential explanation for the current pattern of results is provided by Auer and hinskens [2]. They noted that big cities are typically more ethnically and culturally diverse. Multi-dialectal interactions can speed up dialect leveling, sometimes leading to the formation of regional standard varieties. Applying this idea to Taiwan Mandarin, it is possible that a regional standard variety is currently developing in the big cities and that a more retroflexed (i.e., less merged) sibilant realization is one of its features. A clear prediction following from this hypothesis is that the geographical distribution of merging variation should be sensitive to the age of speakers, with an attenuated effect of urbanization for older speakers as compared to younger speakers. To test this hypothesis, future studies examining geographical effects on sibilant merging for speakers from different age groups are necessary.

5. REFERENCES

[1] Ang, U. 1997. A survey on language use in public occasions in Taiwan, 1997. In: Tung, C., (ed), *Proceedings of the Conference on Language Development in Taiwan*. Hsinchu: Quanmin Bookstore 83–100.

[2] Auer, P., Hinskens, F. 1996. The convergence and divergence of dialects in europe. new and not so new developments in an old area. *Sociolinguistica* 10, 1–30.

[3] Baayen, R. H., Divjak, D. 2017. Ordinal gamms: a new window on human ratings. In: Makarova, A., Dickey, S. M., Divjak, D., (eds), *Each Venture a New Beginning. Studies in Honor of Laura A. Janda*. Bloomington: Slavica 39–56.

[4] Boersma, P., Weenink, D. 2017. *Praat: doing phonetics by computer (version 6.0.22)*.

[5] Chang, Y.-h. S., Shih, C. May 2015. Place contrast enhancement: The case of the alveolar and retroflex sibilant production in two dialects of Mandarin. *Journal of Phonetics* 50, 52–66.

[6] Chuang, Y.-Y. 2009. An Acoustic Study on Voiceless Retroflex and Dental Sibilants in Taiwan Mandarin Spontaneous Speech. Master's thesis National Taiwan University Taipei.

[7] Dong, Z.-s. 1995. First draft of several initial variants in Mandarin under the influence of Chinese dialects in Taiwan. *Yuwen Xuebao* 2, 1–28.

[8] Fasiolo, M., Goude, Y., Nedellec, R., Wood, S. N. 2018. *Fast Calibrated Additive Quantile Regression*. https://arxiv.org/abs/1707.03307v2.

[9] Huang, S. 1993. *Language, society, and ethnic identity – Studies in language sociology in Taiwan]*. Taipei: Crane Publisher.

[10] Huang, Y.-H., Fon, J. 2008. Dialectal variations in tonal register and declination pattern of taiwan mandarin. *Proceedings of Speech Prosody* 605–608.

[11] Ing, R. O. 1984. Issues on the pronunciations of Mandarin. *The World of Chinese Language* 35, 6–16.

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

[13] Jongman, A., Wayland, R., Wong, S. 2000. Acoustic characteristics of English fricatives. *Journal of Acoustical Society of America* 108, 1252–1263. 3.

[14] Kubler, C. C. 1985. The influence of Southern Min on the Mandarin of Taiwan. *Anthropological Linguistics* 27, 156–176. 2.

[15] Labov, W. 1963. The social motivation of a sound change. *Word* 19(3), 273–309.

[16] Li, F., Edwards, J., Beckman, M. E. 2009. Contrast and covert contrast: The phonetic development of voiceless sibilant fricatives in English and Japanese toddlers. *Journal of Phonetics* 37(1), 111–124.

[17] Lin, C. C. 1983. A sociolinguistic study of the use of the retroflex sounds in Mandarin in college studennts in Taiwan. *Bulletin, College of Arts and Letters, National Central University* 1–15.

[18] Rau, D.-h., Li, M.-C. 1994. Phonological variation study of Chinese retroflexed initials (ts) (tsh) (s). *Proceedings of the 4th International Conference on Teaching Chinese as a foreign Language* 345–366.

[19] Sankoff, G. 2002. Linguistic outcomes of language contact. In: Chambers, J., Trudgill, P., N., S.-E., (eds), *The Handbook of Language Variation and Change*. Blackwell 638–668.

[20] Shih, Y.-t. 2012. *Taiwanese-Guoyu bilingual children and adults' sibilant fricative production patterns*. PhD thesis Ohio State University.

[21] Thomason, S. G. 2001. *Language contact*. Edinburgh: Edinburgh University Press.

[22] Trudgill, P. 1974. Linguistic change and diffusion: Description and explanation in sociolinguistic dialect geography. *Language in society* 3(2), 215–246.

[23] Tse, J. K.-P. 1998. Do the young people of Taiwan really not distinguish between zh, ch, sh and z, c, s. *The World of Chinese Language* 90, 1–7.

[24] Wieling, M., Nerbonne, J., Baayen, R. H. 2011. Quantitative social dialectology: Explaining linguistic variation geographically and socially. *PloS one* 6(9), e23613.

[25] Wieling, M., Valls, E., Baayen, R. H., Nerbonne, J. 2018. Border effects among catalan dialects. In: Speelman, D., Heylen, K., Geeraerts, D., (eds), *Mixed-Effects Regression Models in Linguistics*. Springer 71–97.

[26] Wood, S. N. 2004. Stable and efficient multiple smoothing parameter estimation for generalized additive models. *Journal of the American Statistical Association* 99(467), 673–686.

---

[1] The population data was obtained from the online demographic data released by the Department of Household Registration, Ministry of the Interior, Taiwan. (https://www.ris.gov.tw/app/portal/346)