

## **Dialectal Variation of the Effect of Prosodic Prominence on Diphthong Reduction in Taiwan Mandarin – using /aɪ/ as an example**

Chieh-Ching Chen<sup>1</sup> & Janice Fon<sup>2</sup>

<sup>1</sup>*National Taiwan University (Taiwan)*, <sup>2</sup>*National Taiwan University (Taiwan)*  
r10142003@ntu.edu.tw, jfon@ntu.edu.tw

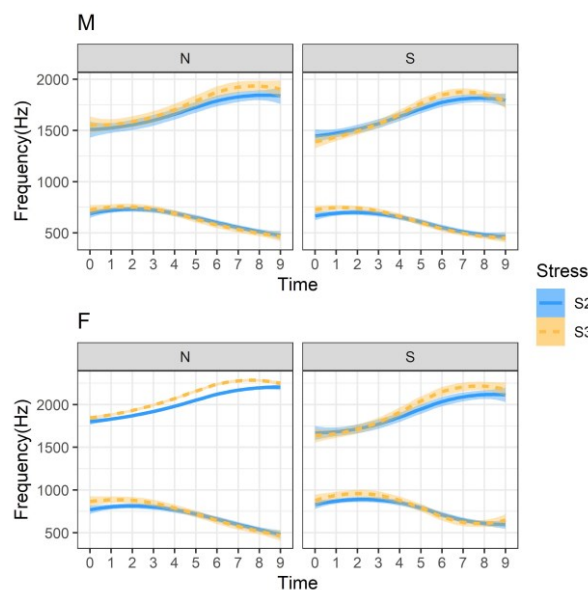
Vowel reduction is frequently observed in spontaneous speech [1]. Previous studies show that prosodic prominence greatly influences the reduction of vowels in English [2], Dutch [3], and Swedish [4]. As prosodic strength is manifested through tonal realization in Mandarin [5], not vowel reduction itself, which is more commonly found in stress-timed languages like English, it would be interesting to see whether and how vowel reduction is affected by stress. Moreover, previous studies on vowel reduction mainly focused on monophthongs, whose reduction was manifested through formant undershoot and shorter duration [2, 3, 4]. For diphthongs, which differ from monophthongs by having an additional vowel target [6], little is said about whether and how reduction should take place. This study thus examines how the reduction of the diphthong /aɪ/ is realized in Taiwan Mandarin and how stress influences the reduction pattern. Besides, since segmental variations exist among Northern and Southern dialects of Taiwan Mandarin [7] and gender difference was observed in vowel reduction [8], we analyzed male and female speakers of the two dialects separately.

Eight hours of monologue recordings of 16 Mandarin-Min bilinguals were chosen from the Taiwan Mandarin-Min Spontaneous Speech Bilingual Corpus [9]. The speakers were evenly divided into two dialectal groups, in which half of them were male, and half were female. We found 5,127 tokens of /aɪ/ in total and labeled their phonetic realizations accordingly. Three levels of stress, S1-S3, were labeled according to Pan-Mandarin Tone and Break Indices (M-ToBI) [5]. S1 is the lowest level of stress and is used when a tone has completely lost its tonal shape. S2 is the next higher level. It is used when the tone still retains its distinctive contour, even though some of its tonal specifications have been lost. S3 is the highest level of stress. It is used to label tones that are realized with a full-fledged contour. In spontaneous speech, S2 is the most common and could be considered the default stress level [10].

Results showed that only 48.8% tokens of /aɪ/ were retained. 46.9% were monophthongized and 4.3% were reduced to other diphthongs. Monophthongization is mainly in the form of merging rather than deletion, as [e] occupied 65.4% of the reduced monophthongs, and 19.3% were centralized as [ə]. Only 15.3% were reduced to [a], in which the second vowel target was dropped. Logistic regression on reduction rate and linear mixed effects models on duration and formant frequencies were performed to examine the effect of stress, gender, and dialect in vowel reduction. Diphthongs with lower levels of stress were found to have a higher reduction rate. We did not find the effect of S1 in terms of duration and spectral results, but S2 is consistently shorter in duration with formant undershoot on both vowel targets compared with S3. Besides, the reduction of diphthongs likely lacks a negative connotation since no gender difference in reduction rate, duration, or spectral results was found [11]. Moreover, speakers of the northern and southern dialects realized [aɪ] differently and adopted different strategies in the formant undershoot. Generally, female southerners had lower and less fronted tongue position for [a] and more fronted [ɪ] in the realization of retained /aɪ/, whereas southern male speakers had higher and less fronted [a] compared with their northern counterparts. As to the realization of prosodic prominence, it was found that in unstressed [aɪ], the tongue position for [a] was higher, while [ɪ] was lower and less fronted in both dialects. Additionally, [a] was less fronted for female northern speakers, but more fronted for male southern speakers when [aɪ] was unstressed.

In this study, we found that diphthongs were more reduced as the two vowel targets merged and the reduction pattern did not differ across genders. Also, we found that diphthongs realized with more reduced tonal shapes tended to have higher reduction rates, shorter duration, and more

centralized spectral results. Finally, dialectal variation was found to be manifested in the spectral results of stressed and unstressed [aɪ].



**Fig.1** Spectral results of retained [aɪ] across S2 and S3 among males (M) and females (F) of the northern (N) and southern (S) dialects.

## References

- [1] Nakamura, M., Iwano, K., & Furui, S. (2008). Differences between acoustic characteristics of spontaneous and read speech and their effects on speech recognition performance. *Computer Speech & Language*, 22(2), 171-184.
- [2] Fourakis, M. (1991). Tempo, stress, and vowel reduction in American English. *The Journal of the Acoustical society of America*, 90(4), 1816-1827.
- [3] Van Bergem, D. R. (1993). Acoustic vowel reduction as a function of sentence accent, word stress, and word class. *Speech communication*, 12(1), 1-23.
- [4] Lindblom, B. (1963). Spectrographic study of vowel reduction. *The journal of the Acoustical society of America*, 35(11), 1773-1781.
- [5] Peng, S. H., Chan, M. K., Tseng, C. Y., Huang, T., Lee, O. J., & Beckman, M. E. (2005). Towards a Pan-Mandarin system for prosodic transcription. *Prosodic typology: The phonology of intonation and phrasing*, 230-270.
- [6] Gussenhoven, C., & Jacobs, H. (2017). *Understanding phonology*. Routledge.
- [7] Fon, J., Hung, J. M., Huang, Y. H., & Hsu, H. J. (2011). Dialectal variations on syllable-final nasal mergers in Taiwan Mandarin. *Language and Linguistics*, 12(2), 273-311.
- [8] Meunier, C., & Espesser, R. (2011). Vowel reduction in conversational speech in French: The role of lexical factors. *Journal of Phonetics*, 39(3), 271-278.
- [9] Fon, J. (2004). A preliminary construction of Taiwan Southern Min spontaneous speech corpus. National Science Council technical report [NSC-92-2411-H-003-050-].
- [10] 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-Fifth International Conference*.
- [11] Labov, W. (1990). The intersection of sex and social class in the course of linguistic change. *Language variation and change*, 2(2), 205-254.