10th International Conference on Speech Prosody 2020
25-28 May 2020, Tokyo, Japan

Phonetic Realisation of Narrow Focus in Wu-Mandarin Bilinguals

Wenxi Fei, Mingyu Weng and Albert Lee

Dept. of Linguistics and Modern Language Studies, The Education U. of Hong Kong, Hong Kong
s1119503@s.eduhk.hk, s1123273@s.eduhk.hk, albertlee@eduhk.hk

Abstract

Many languages have been classified in terms of whether post-focus compression of f_o range (PFC) is used to mark narrow focus or not. While the previous findings may seem to suggest that bilinguals' language dominance could override the contact effect, it is still unclear how language dominance alone affects PFC. Thus, this study tested Suzhou Wu-Mandarin bilinguals (Wu-dominant vs. Mandarin-dominant), who spoke two +PFC languages on a daily basis. We recruited six female Wu-Mandarin bilinguals, four of which self-identified as Mandarin-dominant while two others Suzhou Wu-dominant. Participants were instructed to read aloud questions and corresponding answers in pairs in Suzhou Wu. For narrow focus conditions, the leading question contained one piece of wrong information, which would then elicit contrastive focus in the answer sentence (Initial Focus, Medial Focus, Final Focus). The data from four Mandarin-dominant speakers suggested that PFC was almost lost while the data of two Wu-dominant speakers showed clear PFC in Suzhou Wu speech, no matter which word was under focus. These findings are taken to indicate that bilinguals' language dominance can influence the realisation of PFC in their speech.

Keywords: Post-focus compression, prosodic focus, language dominance

1. Introduction

Post-focus compression (PFC) is a focus-marking strategy where the f_o, duration, and intensity range of post-focus syllables are compressed [1]. In the past decade, many languages have been classified according to whether PFC is used to mark a narrow focus or not [2]. PFC can be found in many Asian languages, including Mandarin [3]-[4], Japanese [5]-[6], Korean [7], Shanghai Wu [8], and Suzhou Wu [9] while absent in others, such as Cantonese [10], and Southern Min [1, 11]. One reason for PFC to considered a typological feature lies in the finding that it appears to be independent of word prosody. For example, Beijing Mandarin and Taiwan Mandarin have almost identical tonal inventory and are mutually intelligible, but the former is +PFC and the latter is -PFC [1]. The latter was hypothesised to have lost PFC through contact with Southern Min (-PFC) in Taiwan, a bilingual society.

On an individual level, language background also appears to influence PFC realisation. In a comparable production study [11], young Southern Min-Mandarin bilinguals showed clear PFC in their Mandarin speech while older speakers did not. As younger speakers in Mainland China are often Mandarin-dominant, vice versa for older speakers, Chen and colleagues' findings may seem to suggest that language dominance is another factor determining PFC realisation. Taken together, [1] and [11] may point to the view that PFC can be lost through contact with a -PFC language, but it may stay intact for speakers whose +PFC language is the dominant one. However, several questions still remain, such as (i) why was the same not observed for Taiwan Mandarin [1], and (ii) whether it follows that two +PFC languages in contact would guarantee bilingual speakers with PFC?

Thus, this study tested Suzhou Wu-Mandarin bilinguals, who spoke two +PFC languages (Suzhou Wu and Mandarin) on a daily basis. Given [1], we hypothesised that all Suzhou Wu-Mandarin bilinguals would show PFC in their Wu speech. If only some speakers show PFC, we may interpret [1]'s findings to mean that PFC is *not* lost through contact with a -PFC language, but through contact with *any* language in a bilingual society. Furthermore, if one group shows clear evidence of PFC while the other does not, we may argue that PFC realisation is closely related to bilinguals' language dominance like in [11].

2. Methodology

2.1. Participants

Six young female bilinguals (age M = 22.17, SD = 1.83) were recruited for a production experiment. They were first asked to complete the Language Background Questionnaires (LBQ) adapted from [12], which has been shown to be valid and reliable [13]. It includes factors such as year of education, the medium of instruction, age of learning (AOL) Mandarin and Suzhou Wu, frequency of language use (FLU), and specific language ability (SLA). Answers to each question in FLU and SLA were on a 1-7 scale. The results showed that all participants spoke Wu at home, Mandarin at school, so they could speak both languages with native competence. Speakers were then categorised into two groups (Wu-dominant vs. Mandarin-dominant) by comparing their self-report of FLU and SLA because other items did not show great differences between the two groups (Table 1). Among the participants, four of them self-identified as Mandarin-dominant while two considered themselves Suzhou Wu-dominant. Note that even the Wu-dominant speakers were highly competent in Mandarin, due to their daily exposure to the language through education, media, and popular culture. As their Mandarin speech may be "at ceiling" and not different from their Mandarin-dominant counterpart, here we first analysed their Wu speech.

2.2. Procedure and stimuli

Participants were recorded in a quiet room using an Audio-Technica AT9934USB microphone (44.1kHz, 16bits). Recording commenced after completing the LBQ. Participants were seated in front of a computer screen, which displayed one question-answer pair at a time. For narrow focus conditions, the leading question contained one piece of wrong information, which would then elicit contrastive focus in the answer sentence (Initial Focus, Medial Focus, Final Focus). For Neutral Focus, the leading question corresponds to the answer sentence. Each

10.21437/SpeechProsody.2020-51

Table 1: *Mean LBQ scores from two groups of Suzhou Wu-Mandarin bilinguals.*

Group	Age	Education (yrs)	AOL (month)		FLU (4-28 scale)		SLA (4-28 scales)	
			MD	SW	MD	SW	MD	SW
Mandarin-dominant	21.8 (1.5)	18.8 (0.5)	26.5 (23.2)	12.8 (7.6)	24 (2.9)	11.5 (3)	26.5 (1.9)	16.8 (4.3)
Suzhou Wu-dominant	23 (2.8)	17.5 (2.1)	36 (17)	6 (0)	17 (0)	18.5 (0.7)	20.5 (0.7)	24 (0)

Note. MD = Mandarin, SW = Suzhou Wu, Mean and Standard Deviation (in parentheses) are shown for the information above.

target sentence contained three words and five syllables, among which the first and third words had two syllables each. The target sentences covered all four lexical tones in Mandarin (MD Tones) for reference (Table 2). When designing the target stimuli, we used Mandarin phonology as a reference for consistency's sake due to the lack of an up-to-date reference of Suzhou Wu phonology and extensive sound change that had taken place since the last systematic description [14]. In total, 80 target sentences were recorded for each speaker (4 tones * 4 focus conditions * 5 repetitions).

Table 2: *Target sentences used in this study.*

MD Tones		Word 1 (sb.)	Word 2 (verb)	Word 3 (sth.)
1	Characters	阿潘	吃	香蕉
	Gloss	"A Pan"	"eat" (s)	"the banana"
2	Characters	徐盈	玩	皮球
	Gloss	"Xu Ying"	"play" (s)	"rubber ball"
3	Characters	小马	打	蚂蚁
	Gloss	"Xiao Ma"	"hit"(s)	"the ants"
4	Characters	万丽	卖	拌面
	Gloss	"Wan Li"	"sell"(s)	"the mixed noodles"

2.3. Data analysis

Data were analysed using Praat (ver. 6.1.06) [15] and ProsodyPro.praat (ver. 5.7.8.2) [16], which extracted time-normalised f_o tracks and duration of each syllable. For each labelled interval, 10 f_o measurements were taken, yielding 50 data points for each utterance. All f_o data were converted into semitones using individual speakers' mean f_o as reference:

$$ST = 12 \, Log_2 (f_o / Mean \, f_o) \quad (1)$$

The resultant f_o data were visualised using Smoothing Spline ANOVA (SS ANOVA) [17]-[18] to examine whether the f_o contours of different focus conditions differed from one another at the different time points. This analysis was conducted in R [19] and the figures were produced using the *ggplot2* package [20]. The shading superimposed on the splines were 95% Bayesian confidence intervals. At any point along the X-axis, if two f_o contours do not overlap, they are considered significantly different from each other.

Moreover, duration and intensity data were compared to corresponding Neutral Focus data and converted into ratios. The ratio data were visualised in boxplots for the comparison of the data in the three focus positions. Then, the raw data were further analysed using two-way ANOVA (Speaker Groups * Focus Positions). Undefined data were deleted in advance.

3. Results

The results showed that for the Suzhou Wu-dominant speakers there was clear evidence of PFC in their Wu speech for both Initial Focus and Medial Focus; for Mandarin-dominant speakers, some PFC was observed only for Initial Focus, but not for Medial Focus.

3.1. f_o contours

Fig. 1 shows the f_o contours of the average f_o contours of target sentences produced by Suzhou Wu-dominant speakers. The post-focus portion of the f_o contours was lower for Initial Focus (lime) and Medial Focus (turquoise), compared with their Neutral Focus counterpart (lilac), i.e. clear evidence of PFC. For Medial Focus and Final Focus, the on-focus f_o peak was higher than Neutral Focus; the same was not observed for Initial Focus. However, the four f_o contours of pre-focus portions do not show great differences. No pre-focus f_o modification was observed.

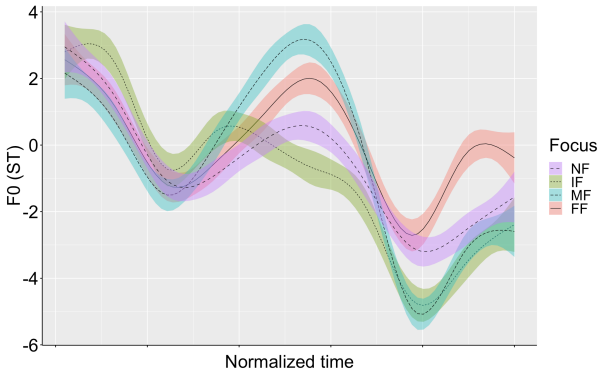


Figure 1: *Time-normalised f_o contours of the Suzhou Wu-dominant group under different focus conditions.*

For the Mandarin-dominant group, PFC was almost absent (Fig. 2). For Initial Focus, post-focus f_o was lower than corresponding Neutral Focus contour for approximately half a syllable; otherwise, the two contours largely overlapped. For Medial Focus, no sign of PFC was observed. There was clear evidence of on-focus f_o peak raising for all focus conditions.

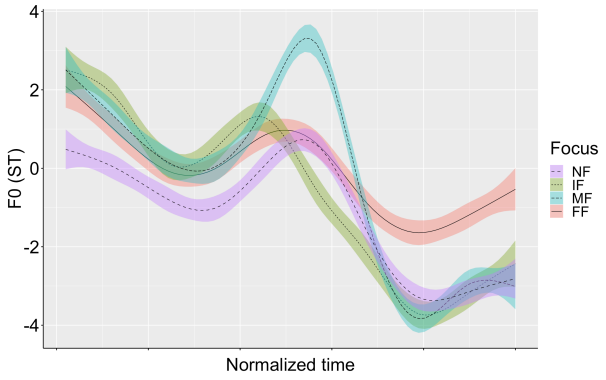
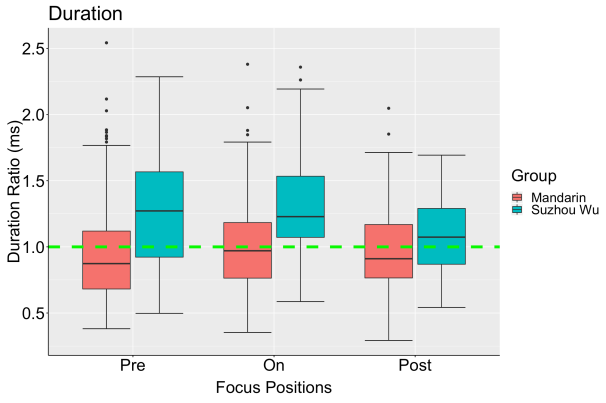


Figure 2: *Time-normalised f_o contours of the Mandarin-dominant group under different focus conditions.*

3.2. Duration and intensity

Fig. 3 illustrates the ratio data of duration and intensity in the three focus positions. Generally, the data of Wu-dominant group were higher than those of Mandarin-dominant one. If we compare the on- and post-focus syllables, we can see that the clear compression of post-focus duration appeared in Suzhou Wu-dominant speakers' speech, but the Mandarin-dominant group's duration did not have great differences between on- and post-focus positions. The intensity data of both groups showed their compression in the post-focus position but it was hard to compare the post-focus data with their on-focus counterparts from this graph. For the pre-focus data, there were no significant contrasts between them and the data of other positions.

As it is mentioned in §2.3, the raw data of duration and intensity have been converted into ratios. If a ratio > 1 (i.e. above the green dashed line), it means the data of that position is higher the value of Neutral Focus sentences (NF); if a ratio < 1, it means that data of that position is lower than NF. Referring to NF, most Wu-dominant data were higher than it. Differently, the data of Mandarin-dominant speakers were close to NF.



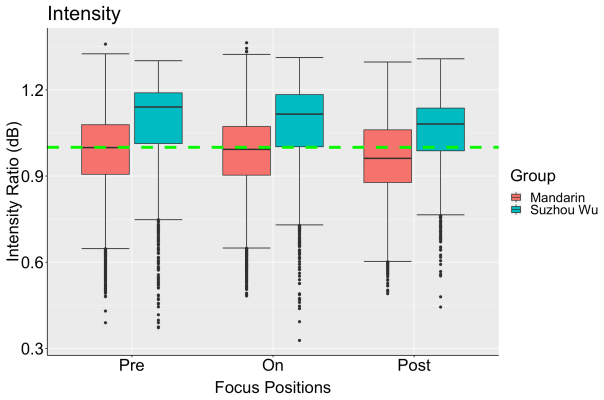


Figure 3: *Duration and intensity ratio of three focus positions between the two groups.*

Table 3 shows the results of duration and intensity using ANOVA. In terms of duration, ANOVA showed that there were significant effects of Speaker Groups, Focus Positions and a significant interaction between them. Post-hoc Tukey test demonstrated that the Wu-dominant and Mandarin-dominant groups significantly differed at $p < .001$ and the difference between the average value of the Wu group and that of the Mandarin group was 65.9 ms. The values of on- and post-focus were more significantly different from each other ($p < .001$). The values of on-focus positions were higher than NF by 27.8 ms. Those of post-focus parts were higher by 0.9 ms and lower than those of on-focus parts by 26.8 ms. For the intensity data, there were significant main effects of Speaker Groups, Focus Positions, and their interaction. By post-hoc Tukey test, the Suzhou Wu-dominant and Mandarin-dominant groups were significantly different at $p < .001$ with the Mean data of the Wu group higher than that of the Mandarin group by 5.9 dB. The values between on- and post-focus were significantly different ($p < .001$). The narrow focus in on-focus positions was higher in intensity than NF by 0.8 dB and that in post-focus ones was lower by 0.5 dB. There was no difference between pre- and on-focus.

Table 3: *ANOVA of duration and intensity.*

	Source	***df***	***F***	***p***
Duration	Residuals	2392	N/A	N/A
	SG	1	284.8	< .001***
	FP	3	13.1	< .001***
	SG*FP	3	9.3	< .001***
Intensity	Residuals	23925	N/A	N/A
	SG	1	2371.6	< .001***
	FP	3	33.0	< .001***
	SG*FP	3	4.0	.007**

Note. SG = Speaker Groups, FP = Focus Positions, SG*FP = Speaker Groups*Focus Positions, *p < .05, **p < .01, ***p < .001.

4. Discussion

The present study shows that some bilinguals can acquire PFC in their Wu speech while others cannot, which might suggest that PFC is "hard-to acquire" even for bilinguals who master two +PFC languages. The results showed PFC in the Suzhou

Wu-dominant group, alongside on-focus raising. For the Mandarin-dominant group, there was on-focus raising but not PFC – in the post-focus regions narrow focus and Neutral Focus contours largely overlapped. Between those two groups, duration and intensity were significantly different as well. Given the findings for the Mandarin-dominant group, our hypothesis that "all Suzhou Wu-Mandarin bilinguals would show PFC in their Wu speech" was not supported.

Unexpectedly, the data of pre-focus intervals were also reported differently in the two groups. In our research, there were pre-focus syllables raising in the Mandarin-dominant group while the four f_o contours of the Wu-dominant group did not show great differences in their pre-focus portions. However, as many researchers have found, pre-focus f_o range and values tend to stay neutral and to be largely individually dependent in +PFC languages [2, 6, 21]. Thus, whether pre-focus syllables may get influenced by language dominance remains to be studied in future.

Our results complement previous studies. This study agrees with Xu et al. that considered language contact as the main factor of the loss of PFC by analysing Beijing Mandarin, Taiwan Mandarin and Taiwanese [1]. It also discusses the possibility that language dominance represented from different age groups could have some effects on PFC than language contact in Chen et al.'s findings [11]. It is further suggested that PFC gets strong effects from the language dominance of bilinguals. To be more specific, when two +PFC languages are brought into contact, PFC will get lost in one of the groups, which means language dominance may play a more important role in this study than contact effects reported by Xu et al. [1]. Moreover, from the typological perspective, the results of the Wu-dominant group are also consistent with Wang, Zhang, Xu and Ding's study that Suzhou Wu is a +PFC language [9].

Compared with [22], our study may explain the results in a similar way, which argues that language dominance, including frequency of language use (FLU), specific language ability (SLA), can bring about the failure of performing PFC. Liu, Xu and Lee demonstrated that PFC is closely associated with bilinguals' language proficiency. They found that the advanced speakers whose L1 were Korean had some evidence of PFC in their English speech (L2), while beginners and intermediate learners did not show PFC in their L2 production [22]. However, the data of advanced speakers were largely far away from that of native English speakers. On the one hand, its results indicated insufficient language proficiency might lead to the negative transfer of PFC from Korean (+PFC) to English (+PFC). On the other hand, the positions of post-focus f_o could still showed the difference between non-native speakers and native speakers, which suggested that the realisation of f_o can tell whether a speaker is native-like.

In the present study, it was noticed that Wu-dominant speakers mastered both Mandarin and Suzhou Wu well with a slightly higher preference for Suzhou Wu while Mandarin-dominant participants strongly favoured using Mandarin in their daily life. Referring to LBQ (Table 1), the scores of the Wu-dominant group in FLU and SLA of speaking Wu were both higher than those of speaking Mandarin, but the differences were not great. The scores of the two languages can be shown by absolute values: |MD-SW| = 12.5 (FLU); 9.7 (SLA) in the Mandarin-dominant group while |MD-SW| = 1.5 (FLU); 3.5 (SLA) in the Wu-dominant group (MD = Mandarin, SW = Suzhou Wu). In summary, Wu-dominant speakers relatively achieved a balance of the proficiency of the two languages whereas Mandarin-dominant speakers were generally not quite confident in their proficiency in Suzhou Wu.

Different from [22], our data show that both groups had high proficiency in Mandarin and saw Suzhou Wu as their mother tongue, so it is less likely to explain their languages as L1 or L2. Therefore, we suggest that for bilinguals who are native speakers of both +PFC languages, their self-estimated proficiency can be further discussed and may lead to their unsuccessful production of PFC. Additionally, the above differences are possible to be attributed to their AOL of the two languages and family, education or work effects reported by their LBQ but the education years and ages of test do not show great differences between different participants.

This paper unfolds a new perspective on the connection between bilingualism and language production after some scholars noticed that language dominance can influence language production [23]. Besides, with the updated recordings from young adults, our data could also be used to track the sound change in time and to add more descriptions on the phonological system of Suzhou Wu and Mandarin. However, the scale of our study might be not large enough to reach a solid conclusion. Thus, more participants could be recruited and more data of LBQ can be analysed in detail.

5. Conclusions

In conclusion, this paper is consistent with that the prosodic focus of bilinguals may be affected by their language dominance. It indicated that Suzhou Wu-dominant speakers compressed their post-focus words clearly while Mandarin-dominant speakers do not have significant changes in their post-focus portions compared with the counterparts of Neutral Focus sentences. Interestingly, we noticed that f_o contours in the pre-focus parts are performed differently in the two groups as well. Lastly, this study provides the up-to-date prosodic data of the young people in Suzhou Wu and connects PFC, a typological marker, with bilingualism. Nevertheless, it also remains to be further analysed with larger-scaled investigations.

6. References

[1] Y. Xu, S.-W. Chen, and B. Wang, "Prosodic focus with and without post-focus compression: A typological divide within the same language family?," *The Linguistic Review*, vol. 29, no. 1, pp. 131-147, 2012.

[2] Y. Xu, "Post-focus Compression: Cross-linguistic Distribution and Historical Origin," in *Proceedings of the 17th International Congress of Phonetic Sciences (ICPhS 2011)*, 2011, pp. 152-155.

[3] Y. Chen and C. Gussenhoven, "Emphasis and tonal implementation in Standard Chinese," *Journal of Phonetics*, vol. 36, no. 4, pp. 724–746, 2008.

[4] F. Liu and Y. Xu, "Parallel Encoding of Focus and Interrogative Meaning in Mandarin Intonation," *Phonetica*, vol. 62, no. 2-4, pp. 70–87, 2005.

[5] S. Ishihara, "Syntax-phonology interface of wh-constructions in Japanese," in *Proceedings of The Coaching Psychologist*, 2002, pp. 165-189.

[6] A. Lee and Y. Xu, "Revisiting focus prosody in Japanese," in *Proceedings of the 6th International Conference on Speech Prosody (SP2012)*, 2012, pp. 274–277.

[7] Y. C. Lee and Y. Xu, "Phonetic realization of contrastive focus in Korean," In *Proceedings of the 5th International Conference on Speech Prosody (SP2010)*, 2010.

[8] Y. Chen, "The acoustic realization of vowels of Shanghai Chinese," *Journal of Phonetics*, vol. 36, no. 4, pp. 629–648, 2008.

[9] B. Wang, Y. Zhang, Y. Xu, and H. Ding, "Prosodic focus in three northern Wu dialects: Wuxi, Suzhou and Ningbo," in *Proceedings*

of 8th ExLing 2017, August, Heraklion, Greece, 2017, pp. 117–120.
[10] W. L. Wu and Y. Xu, "Prosodic focus in Hong Kong Cantonese without post-focus compression," in *Proceedings of the 5th International Conference on Speech Prosody (SP2010)*, 2010.
[11] Y. Chen, Y. Xu, and S. Guion-Anderson, "Prosodic Realization of Focus in Bilingual Production of Southern Min and Mandarin," *Phonetica*, vol. 71, no. 4, pp. 249–270, 2014.
[12] J. E. Flege, M. J. Munro, and I. R. A. Mackay, "Factors affecting strength of perceived foreign accent in a second language," *The Journal of the Acoustical Society of America*, vol. 97, no. 5, pp. 3125–3134, 1995.
[13] D. E. Metz, F. Caccamise, and M. S. Gustafson, "Criterion validity of the Language Background Questionnaire: A self-assessment instrument," *Journal of Communication Disorders*, vol. 30, no. 1, pp. 23–32, 1997.
[14] X. Ye and Y. Sheng, *The audio collection of Suzhouhua (Modern Chinese Dialect Sound Database) (J. Hou, Ed.)*. Shanghai: Shanghai Educational Publishing House, 1996.
[15] P. Boersma and D. Weenink, *Praat, a system for doing phonetics by computer*. Computer program, Version 6.1.06. Available: http://www.fon.hum.uva.nl/praat/. [Accessed: 18-Dec-2019].
[16] Y. Xu, "ProsodyPro — A Tool for Large-scale Systematic Prosody Analysis," in *Proceedings of Tools and Resources for the Analysis of Speech Prosody (TRASP 2013), Aix-en-Provence, France*, 2013, pp. 7-10.
[17] L. Davidson, "Comparing tongue shapes from ultrasound imaging using smoothing spline analysis of variance," *The Journal of the Acoustical Society of America*, vol. 120, no. 1, pp. 407–415, 2006.
[18] C. Gu, "Smoothing spline ANOVA models: R package gss," *Journal of Statistical Software*, vol. 58, no. 5, pp. 1-25, 2014.
[19] R Core Team (2019). *R: A language and environment for statistical computing (version 3.6.1)*. Available: http://www.R-project.org/ [Accessed: 18-Dec-2019].
[20] H. Wickham, *ggplot2: elegant graphics for data analysis*, Springer New York, 2009.
[21] Y. Xu, C. X. Xu and X. Sun, "On the temporal domain of focus," in *Proceedings of the 2nd International Conference on Speech Prosody (SP2004)*, 2004.
[22] J. Liu, Y. Xu, and Y.-C. Lee, "Post-focus compression is not automatically transferred from Korean to L2 English," *Phonetics and Speech Sciences*, vol. 11, no. 2, pp. 15–21, 2019.
[23] G. T. Lourido and B. G. Evans, "The effects of language dominance switch in bilinguals: Galician new speakers speech production and perception," *Bilingualism: Language and Cognition*, vol. 22, no. 3, pp. 637–654, 2018.