EFFECTS OF FOCUS ON DURATION AND INTENSITY IN CHONGMING CHINESE

Yike Yang[1], Si Chen[1] and Kechun Li[2]

[1]Department of Chinese and Bilingual Studies, The Hong Kong Polytechnic University
[2]Department of Theoretical and Applied Linguistics, University of Cambridge
yi-ke.yang@connect.polyu.hk, sarah.chen@polyu.edu.hk, kl502@cam.ac.uk

ABSTRACT

Prosodic focus is generally realised by expanded pitch range, lengthened duration and increased intensity on the focused components, while the post-focus components may be associated with a compressed pitch range and intensity, which is referred to as post-focus compression (PFC). However, controversy surrounds whether PFC exists cross-linguistically, and detailed studies on how focus influences duration and intensity are scarce. This study aims to contribute novel data to the prosodic typology literature by examining the effects of focus on duration and intensity in Chongming Chinese (CC). Twenty target words were embedded under different focus conditions, and the production data were submitted for linear-mixed effects models. Our results showed focus-induced change of duration and intensity (i.e., lengthened duration and a larger intensity range under focus) as well as PFC of duration and intensity range in CC.

Keywords: focus, prosody, duration, intensity, Chongming Chinese

1. INTRODUCTION

Prosodic focus highlights part of an utterance and is generally realised by expanded pitch range, lengthened duration and increased intensity on the focused components [1]. The post-focus components may be associated with reduced or compressed pitch range [2], [3], the phenomenon of which has been coined as post-focus compression (PFC) [4]. However, the presence or absence of PFC seems to vary even within the same language family [5].

Another issue worthy of examination is the potential interaction between prosodic focus and lexical tones in tone languages. Tone languages use pitch to distinguish lexical items, which may interact with or even prevent the realisation of prosodic focus. According to [6], focus is not expressed by prosodic means in Yucatec Maya, which the authors attributed to the fact that pitch is already used for marking tones. Meanwhile, [7] and [8] reported the effect of focus on the pitch curves of local tone-bearing units, revealing the influences of both tone and focus on pitch realisation. However, little is known about the effects of focus on duration and intensity. Following this line of research, the current project aims to investigate the realisation of prosodic focus in a tone language, Chongming Chinese (CC), and attempts to examine the interaction between focus and tone and their influence on duration and intensity.

CC is an under-documented Chinese dialect that is mainly spoken in Chongming County, Qidong City and Haimen City in Eastern China. Because possible variations of tonal representation among different age groups have been indicated [9], we examined the older version of this dialect and invited only middle-aged speakers. As shown in Table 1 (adapted from [10]), CC has eight lexical tones, of which two are level tones (Tones 1 and 5), four are contour tones (Tones 2, 3, 4 and 6) and the remaining two are checked tones (Tones 7 and 8). Adopting the five-point tone scale of [11], we used 5 to represent the highest pitch level and 1 for the lowest.

Table 1: Tone system in CC.

Tone	1	2	3	4	5	6	7	8
Pitch value	55	24	424	242	33	313	5	2

Based on the gaps in the field, the current paper is an attempt to address the following research questions:

1) What are the effects of focus and tone on duration and intensity in CC?

2) Is PFC present or absent in CC? If present, how is it realised?

2. METHOD

2.1. Participants

Twelve native speakers of CC (six females; six males) aged between 38 and 57 (mean ± SD: 52.00 ± 4.53) were recruited. They self-reported that CC is their native and dominant language as well as the language they use for daily communication. No participants have received formal musical training, and none reported any history of speaking, hearing or language difficulty.

2.2. Stimuli and procedure

Twenty monosyllabic words varying in tones were selected as the target stimuli. Because [10] observed that the vowel [æ] bears the largest range of tones for different onsets in CC, we included twelve words containing [æ] and further added another eight words. We also considered the context effect and selected two preceding and following syllables respectively. The stimuli were then embedded in carrier phrases with different contextual combinations.

We manipulated four focus conditions: neutral, initial, medial and final foci. For the neutral focus condition, the participants were instructed to read the target sentence at a normal speech rate. For the remaining three conditions, precursor questions were asked to elicit different focus locations. The participants then answered the questions using the texts shown on a computer screen. The focus was located on the preceding syllable for the initial focus condition, on the target syllable for the medial focus condition and on the following syllable for the final focus condition. In total, there were 3,840 sentences (20 target syllables * 4 contexts * 4 foci * 12 speakers).

The recording took place in a quiet room in Qidong City. The stimuli were randomly presented in E-Prime [12] on one PC, and all the utterances were recorded with Praat [13] on another PC.

2.3. Data analysis

The sonorants of the preceding, target and following syllables were manually segmented by trained phoneticians using Praat. The criteria for segmentation followed the conventions described in [14]. After the segmentation, duration and intensity values were extracted using the ProsodyPro Praat script [15].

To test the focus effect and PFC, the medial and initial foci (each corresponding to focused and post-focused syllables) were compared with the neutral focus (unfocused syllables). The final focus was not included in the analysis. We first examined how duration was affected by focus, tone and context with linear mixed effects models [16], [17] in R [18]. Next, we compared how intensity contour and range were affected by these variables. The figures were plotted using the 'ggplot2' [19] package.

3. RESULTS

3.1. Duration

Main effects of focus ($t[2] = 36.576$, $p < .001$), tone ($t[7] = 728.04$, $p < .001$) and preceding syllable ($t[1] = 27.629$, $p < .001$) were found for the duration of the target syllables. No two-way interactions reached significance. The unfocused syllables were 12.3 ms shorter than the focused ones ($p < .001$) and 9.6 ms longer than the post-focused ones ($p = .003$).

We then analysed each tone individually to prevent any net effect [20]. Main effect of focus and preceding syllable was significant for most of the tones. The average duration values for each tone and focus condition are plotted in Figure 1.

Figure 1: Duration by tone and focus.

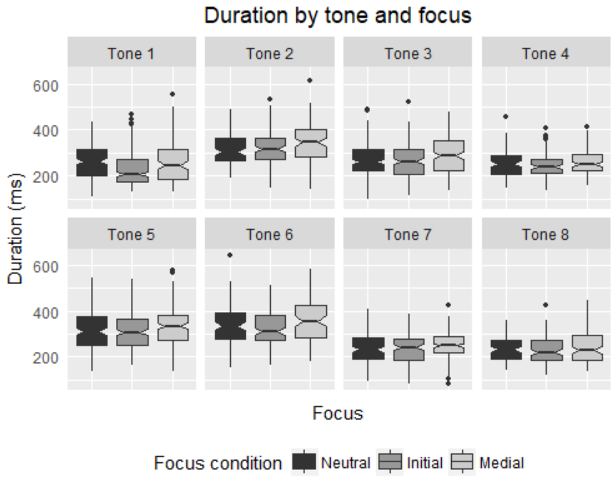


Table 2: Models comparing duration of neutral and medial foci (neutral as the baseline).

Tone	Estimate	SE	t value	Pr(>|t|)
Tone 1	0.2395	9.4466	0.025	0.980
Tone 2	21.51	10.71	2.008	0.0462*
Tone 3	16.041	7.802	2.056	0.0407*
Tone 4	7.656	7.439	1.029	0.3049
Tone 5	16.173	8.899	1.817	0.0702
Tone 6	18.63	12.52	1.489	0.138
Tone 7	15.269	6.364	2.399	0.0171*
Tone 8	5.067	8.472	0.598	0.55

Table 3: Models comparing duration of neutral and initial foci (neutral as the baseline).

Tone	Estimate	SE	t value	Pr(>|t|)
Tone 1	-30.359	8.880	-3.419	0.0007***
Tone 2	-0.0291	9.4487	-0.003	0.9975
Tone 3	-4.964	7.817	-0.635	0.526
Tone 4	-10.25	7.04	-1.456	0.1472
Tone 5	-5.111	8.501	-0.601	0.548
Tone 6	-15.97	10.98	-1.455	0.147
Tone 7	-2.904	6.328	-0.459	0.647
Tone 8	-5.666	7.478	-0.758	0.4497

We then fitted models for each tone and focus pair (initial and medial foci with the neutral focus) to examine the focus effects in detail. In general, the focused syllables had longer duration than their unfocused counterparts as shown in Table 2, while the unfocused ones were longer than the post-focused syllables, as in Table 3.

Our data showed consistent lengthening of the focused syllables across tones as well as decreased duration of post-focused syllables in CC.

3.2. Intensity

For intensity values, we first analysed the effects of the variables on the 20 normalised time points. Significant main effects of focus ($t[2] = 4484.2$, $p < .001$), tone ($t[7] = 954.74$, $p < .001$), preceding syllable ($t[1] = 7.928$, $p = .005$) and following syllable ($t[1] = 22.02$, $p < .001$) were found. There was a significant two-way interaction between focus and tone ($t[14] = 117.17$, $p < .001$). Both focused and post-focused syllables were larger in intensity than the unfocused ones ($p < .001$ for both).

Again, each tone was analysed individually. Focus significantly influenced the intensity values of all tones. There was a significant main effect of the following syllable for three tones and a significant main effect of the preceding syllable for only one tone. The intensity contours are shown in Figure 2.

Figure 2: Intensity contours by tone and focus.

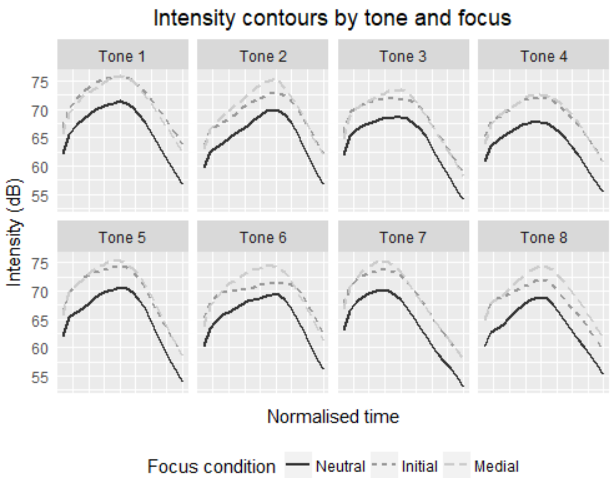


Table 4: Models comparing intensity contours of neutral and medial foci (neutral as the baseline).

Tone	Estimate	SE	t value	Pr(>|t|)
Tone 1	4.70212	0.19440	24.19	< 2e-16***
Tone 2	5.02184	0.28110	17.865	< 2e-16***
Tone 3	3.94376	0.21981	17.94	< 2e-16***
Tone 4	4.97827	0.27739	17.95	< 2e-16***
Tone 5	4.3995	0.2083	21.12	< 2e-16***
Tone 6	4.21083	0.27651	15.228	< 2e-16***
Tone 7	4.88643	0.21607	22.61	< 2e-16***
Tone 8	5.2463	0.2648	19.81	< 2e-16***

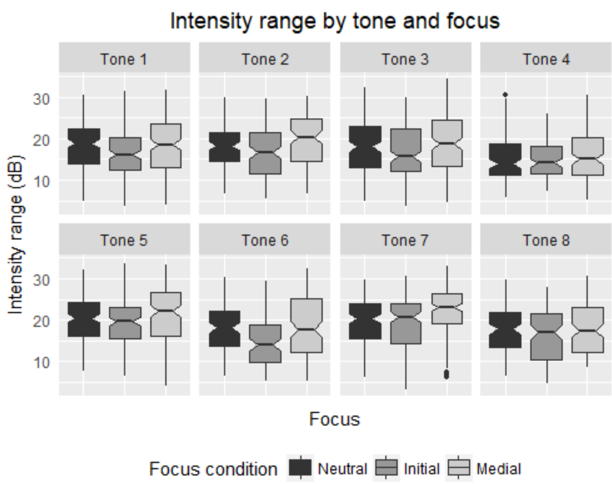
Table 5: Models comparing intensity contours of neutral and initial foci (neutral as the baseline).

Tone	Estimate	SE	t value	Pr(>|t|)
Tone 1	5.06633	0.20186	25.10	< 2e-16***
Tone 2	4.15122	0.25920	16.015	< 2e-16***
Tone 3	4.04182	0.19324	20.92	< 2e-16***
Tone 4	4.74471	0.26302	18.04	< 2e-16***
Tone 5	4.25464	0.20452	20.80	< 2e-16***
Tone 6	3.56077	0.23944	14.871	< 2e-16***
Tone 7	3.75645	0.19452	19.31	< 2e-16***
Tone 8	3.57443	0.26343	13.57	< 2e-16***

We also fitted models for intensity under each tone and focus pair. The focused syllables had significantly lower intensity than their unfocused and post-focused counterparts under all tones, as listed in Tables 4 and 5.

Next, we calculated the intensity range of each syllable and fitted new models with the range as the dependent variable. Significant main effects of focus ($t[2] = 70.158$, $p < .001$), tone ($t[7] = 183.72$, $p < .001$), preceding syllable ($t[1] = 38.46$, $p < .001$) and following syllable ($t[1] = 19.402$, $p < .001$) were found. The unfocused syllables had a smaller intensity range than the focused ones ($p < .001$) but a larger intensity range than the post-focused ones ($p < .001$), although the differences were small (roughly 1 dB difference for each pair). The intensity range by tone and focus is plotted in Figure 3.

Figure 3: Intensity range by tone and focus.



The models comparing the intensity range of different focus pairs are listed in Tables 6 and 7, which reveal a general trend of the focused syllables exhibiting a larger intensity range than their unfocused counterparts and the unfocused syllables showing a larger range than the post-focused ones.

Table 6: Models comparing intensity range of neutral and medial foci (neutral as the baseline).

Tone	Estimate	SE	t value	Pr(>|t|)
Tone 1	-0.04005	0.68995	-0.058	0.954
Tone 2	1.5316	0.7602	2.015	0.0454*
Tone 3	1.1029	0.6884	1.602	0.11
Tone 4	0.8703	0.7595	1.146	0.253
Tone 5	1.3304	0.6494	2.049	0.0414*
Tone 6	0.6251	0.9192	0.68	0.497
Tone 7	1.8671	0.6009	3.107	0.00209**
Tone 8	0.3369	0.8135	0.414	0.679

Table 7: Models comparing intensity range of neutral and initial foci (neutral as the baseline).

Tone	Estimate	SE	t value	Pr(>|t|)
Tone 1	-2.2625	0.6093	-3.713	0.00025***
Tone 2	-1.1360	0.7945	-1.43	0.155
Tone 3	-0.8702	0.6782	-1.283	0.201
Tone 4	-0.5336	0.7047	-0.757	0.45
Tone 5	-0.5130	0.5999	-0.855	0.393
Tone 6	-3.2363	0.8037	-4.027	8.4e-05***
Tone 7	-0.7216	0.6023	-1.198	0.232
Tone 8	-1.2475	0.7776	-1.604	0.11

4. DISCUSSION

In this study, we investigated the effects of focus and tone on duration and intensity in CC, an under-studied tone language. Main effects of focus and tone on duration and intensity (both contour and range) were found, and a two-way interaction between focus and tone for intensity contour was evident. An earlier study showed that the focused syllables in CC have expanded pitch contour than their unfocused counterparts [8], and our data proved that the focused syllables also exhibit lengthened duration, higher intensity and a larger intensity range. Taken together, these results suggest that, typologically, CC belongs to the languages that have typical realisation of prosodic focus [21], [22].

In terms of PFC, the duration values were reduced in the post-focus position, although the effect of such reduction was not robust. In fact, only the overall reduction of all tones and that of Tone 1 were statistically significant. A similar pattern was found for the intensity range, where only the compression of intensity range for all the syllables and Tones 1 and 6 reached significance. Although we provided some evidence for PFC [4], [5] using novel data from CC, it seems that PFC is especially prominent for the high level Tone 1 in CC. This echoes the findings in [23], where PFC occurred exclusively after a pitch accent in Japanese, and the authors postulated that PFC may be realised conditionally depending on language-specific features. Subsequent studies may design more thorough tests to investigate how exactly PFC is realised and conditioned in CC.

Although we have demonstrated the [+PFC] feature of CC, a noteworthy point is that the post-focused syllables were above the unfocused syllables in intensity contour as shown in Figure 2. This is strikingly different from many languages (e.g., Korean [24] and Persian [25]), wherein post-focused syllables all exhibit decreased intensity. This divergence might be the result of different focus domains [26] or syntactic structures [27]. The focus domain of a language may be a syllable, a word or a phrase. If the focus domain of CC is a phrase, and in our case, this phrase includes the preceding syllable as a verb and the target syllable as its object, then the target syllable is still within the focus domain even if it is under the post-focus condition. Another possible explanation has to do with the syntactic structure of our stimuli. Unlike most of the studies on prosodic focus (e.g., [28]), the target syllable in our study was the object of the preceding syllable. NP and VP foci are usually realised differently [27], and recent evidence suggests that the patterns of focus realisation may vary even within complex NPs [29], [30]. It is thus not surprising that our data do not conform to previous findings. If either the focus domain effect or the syntactic structure effect is true, we can also explain why the post-focus reduction of duration is not robust in our data. Again, further investigations are required to test these hypotheses.

Another interesting point is the possible interaction between tone and duration in focus realisation. Earlier work on intonation language revealed typological differences in the interaction between accented syllables and duration when marking focus [31]. Because lexical tone also influences syllable duration, it is necessary for future studies to confirm whether and how lexical tone and duration interact and how they affect focus realisation in CC.

In conclusion, there is focus-induced change of duration and intensity as well as PFC of duration and intensity range in CC. However, the exact conditions governing PFC realisation in CC remain unknown. Attention should also be paid to the scope of focus domain and syntactic structure to gain a better understanding of prosodic focus.

5. ACKNOWLEDGEMENTS

This study was funded by a research grant from Faculty of Humanities, the Hong Kong Polytechnic University (grant number: 1-ZVHJ). We thank all the informants for their participation and the three anonymous reviewers for their constructive suggestions.

6. REFERENCES

[1] Cooper, W. E., Eady, S. J., Mueller, P. R. 1985. Acoustical aspects of contrastive stress in question-answer contexts. *J. Acoust. Soc. Am.* 77(6), 2142–2155.

[2] Ishihara, S. 2002. Syntax-Phonology Interface of Wh-Constructions in Japanese. *Proc. 3rd TCP* Tokyo, 165–189.

[3] Xu Y., Xu, C. X. 2005. Phonetic realization of focus in English declarative intonation. *J. Phon.* 33(2), 159–197.

[4] Xu, Y. 2011. Post-focus compression: Cross-linguistic distribution and historical origin. *Proc.17th ICPhS* Hong Kong, 152–155.

[5] Xu, Y., Chen, S. W., Wang, B. 2012. Prosodic focus with and without post-focus compression: A typological divide within the same language family? *Linguist. Rev.* 29(1), 131–147.

[6] Kügler, F., Skopeteas, S. 2007. On the universality of prosodic reflexes of contrast: The case of Yucatec Maya. *Proc. 16th ICPhS* Saarbrücken, 1025–1028.

[7] Xu, Y. 1999. Effects of tone and focus on the formation and alignment of f0 contours. *J. Phon.* 27(1), 55–105.

[8] Yang, Y., Chen, S., Li, K. 2018. Pitch realization of post-focus components in Chongming Chinese. *J. Acoust. Soc. Am.* 144(3), 1938–1938.

[9] Chen S., Wiltshire, C. 2013. Tone realization in younger vs. older speakers of Nanjing Dialect. In: Zhuo, J.-S. (ed), *Increased Empiricism: Recent advances in Chinese Linguistics*. Amsterdam: John Benjamins, 147–170.

[10] Chen, S. 2014. A Phonetic and Phonological Investigation of the Tone System of Chongming Chinese. University of Florida.

[11] Chao, Y. R. 1930. A system of tone letters. *Le maitre Phon.* 45, 24–27.

[12] Schneider, W., Eschman, A., Zuccolotto, A. 2012. *E-Prime User's Guide*. Pittsburgh: Psychological Software Tools Inc.

[13] Boersma, P., Weenink, D. 2001. Praat, a system for doing phonetics by computer. *Glot international* 5, 341–345.

[14] Zhang, Y., Nissen, S. L., Francis, A. L. 2008. Acoustic characteristics of English lexical stress produced by native Mandarin speakers. *J. Acoust. Soc. Am.* 123(6), 4498–4513.

[15] Xu, Y. 2013. ProsodyPro - A tool for large-scale systematic prosody analysis. *Proc. TRASP 2013* Aix-en-Provence, 7–10.

[16] Baayen, R. H. 2008. *Analyzing linguistic data: A practical introduction to statistics using R*. New York: Cambridge University Press.

[17] Bates, D., Mächler, M., Bolker, B., Walker, S. 2015. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67(1), 1–48.

[18] R Core Team. 2018. R: A Language and Environment for Statistical Computing. Vienna: R Foundation for Statistical Computing.

[19] Wickham, H. 2016. *ggplot2: Elegant Graphics for Data Analysis*. Cham: Springer.

[20] DiCanio, C., Benn, J., García, R. C. 2018. The phonetics of information structure in Yoloxóchitl Mixtec. *J. Phon.* 68, 50–68.

[21] de Jong, K. 2004. Stress, lexical focus, and segmental focus in English: Patterns of variation in vowel duration. *J. Phon.* 32(4), 493–516.

[22] Heldner, M. 2003. On the reliability of overall intensity and spectral emphasis as acoustic correlates of focal accents in Swedish. *J. Phon.* 31(1), 39–62.

[23] Lee, A., Xu, Y. 2018. Conditional realisation of post-focus compression in Japanese. *Proc. Speech Prosody 2018* Poznań, 216–219.

[24] Lee, Y. C., Xu, Y. 2010. Phonetic Realization of Contrastive Focus in Korean. *Proc. Speech Prosody 2010* Chicago, 1–4.

[25] Taheri-Ardali, M., Xu, Y. 2012. Phonetic Realization of Prosodic Focus in Persian. *Proc. Speech Prosody 2012* Shanghai, 326–329.

[26] van Heuven, V. J. 1994. What is the smallest prosodic domain? In Keating, P. A. (ed), *Papers in Laboratory Phonology*. Cambridge: Cambridge University Press, 76–98.

[27] Jun, S.-A. 2011. Prosodic Markings of Complex NP Focus , Syntax , and the Pre-/Post-focus. *Proc. 28th WCCFL* Los Angeles, 214–230.

[28] Chen, Y., Xu, Y., Guion-Anderson, S. 2014. Prosodic Realization of Focus in Bilingual Production of Southern Min and Mandarin. *Phonetica*, 71(4), 249–270.

[29] Hsu Y.-Y., Xu, A. 2017. Focus Acoustics in Mandarin Nominals. *Proc. Interspeech 2017* Stockholm, 3231–3235.

[30] Hsu Y.-Y., Xu, A., Ngai, H. 2018. Focus Prosody in Cantonese and Teochew Noun Phrases. *Proc. Speech Prosody 2018* Poznań, 961–965.

[31] Vanrell, M., Stella, A., Fivela, B., Prieto, P. 2013. Prosodic manifestations of the Effort Code in Catalan, Italian and Spanish contrastive focus. *J. Int. Phon. Assoc.* 43(2), 195–220.