Prosodic Realization of Focus in Changchun Mandarin and Nanjing Mandarin

Ying Chen*, Jiajing Zhang, Bingying Ye and Chenfang Zhou
School of Foreign Studies, Nanjing University of Science and Technology
Nanjing, China
*E-mail: ychen@njust.edu.cn
E-mail: iris_zhang@njust.edu.cn
E-mail: yby2006tlrx@126.com
E-mail: zhouchenfang_19@163.com

Abstract— **This study was designed to explore the prosodic patterns of focus in two dialects of Mandarin. One is Changchun Mandarin and the other is Nanjing Mandarin. The current paper compares the acoustics of their prosodic realization of focus in a production experiment. Similar to standard Mandarin, which uses in-focus expansion and concomitantly post-focus compression (PFC) to code focus, results in the current study indicate that both Changchun and Nanjing speakers produced significant in-focus expansion of pitch, intensity and duration and PFC of pitch and intensity in their Mandarin dialects. Meanwhile, the results show no significant difference of prosodic changes between Changchun and Nanjing Mandarin productions. These results reveal that PFC not only exists in standard Mandarin but also in Mandarin dialects.**

Keywords-prosodic focus; post-focus compression; Changchun Mandarin; Nanjing Mandarin

I. INTRODUCTION

Prosody refers to the suprasegmental features in an utterance. Its acoustic parameters include duration, pitch (F0), intensity, etc. Focus is used to highlight a certain part of an utterance in a certain context, which in many cases, in addition to morphosyntactic means, can be realized by means of prosody [24]. Therefore, prosodic change plays an essential role in focus realization. Reference [21] proposed main features of prosodic change: an increase in F0, intensity, and duration in focused components, a decrease of F0 and intensity in post-focus components (post-focus compression), and no consistent prosodic change in pre-focus components. Post-focus compression (PFC) was found in many languages, including standard Mandarin [21] and English [20], whereas PFC does not exist in some languages, such as Southern Min [11, 24] and Cantonese [14].

References [16, 21, 22] found that standard Mandarin has PFC. Both Changchun Mandarin and Nanjing Mandarin are subdialects of Mandarin Chinese. The former belongs to Northeast Mandarin and the latter belongs to Jianghuai Mandarin. The two dialects are both tonal but with different tonal inventories. Standard Mandarin has four tones: T1 (55), T2 (35), T3 (214) and T4 (51). The numbers in parentheses represent tone values based on Chao's five-scale notation system [19]. The corresponding tones in Changchun Mandarin are T1 (44), T2 (24), T3 (214) and T4 (53) [25], and that in Nanjing Mandarin are T1 (31), T2 (13), T3 (212) and T4 (44) [26].

Based on previous findings in the research of prosodic focus, this study investigates whether PFC exists in these two Mandarin dialects as it exists in standard Mandarin and how native speakers of these Mandarin dialects realize prosodic focus. Three research questions are addressed:

(1) Do Changchun and Nanjing Mandarin speakers produce significant in-focus expansion of duration, F0 and intensity in their dialects?

(2) Do they produce significant post-focus compression (PFC) of F0 and intensity in their dialects?

(3) Do the prosodic focus patterns in Changchun and Nanjing Mandarin production differ from one another?

II. METHODS

A. Participants

Ten Changchun (CC) Mandarin speakers and ten Nanjing (NJ) Mandarin speakers (five males and five females in each group) were recruited from Nanjing University of Science and Technology in China. They are all undergraduate students, age from 18 to 24, born and raised in either Changchun or Nanjing. They speak CC or NJ Mandarin as their first language.

B. Stimuli

The stimuli are listed in Table I and Table II. Participants were instructed to answer the same target sentences in their dialect, guided by different prompt questions. Previous studies found that PFC can be clearly observed in a sentence containing all sonorant segments given that all sonorants are tone-bearing unit [21, 24]. Therefore, each target sentence in this study uses three words and five syllables/characters with sonorant onsets which bear the same tone in standard Mandarin while corresponding to a different tone in CC and NJ Mandarins. Each syllable in the same target sentence in the same dialect is pronounced in the same tone. For instance, the Mandarin tone in the first sentence is Tone 1 but pronounced with tones in different dialects as shown in Table II. There are four target sentences in four tones corresponding to the T1 (55), T2 (35), T3 (214) and T4 (51) in standard Mandarin respectively. To obtain prosodic change of focus, different prompt questions were raised to elicit focus in different locations: initial focus (the subject), medial focus (the verb) and final focus (the object).

TABLE I. PROMPT QUESTIONS AND ANSWERS IN THE STIMULI

		Sentence 1	Sentence 2	Sentence 3	Sentence 4
Neutral focus	Q	ni shuo shenme? *'What did you say?'*	ni shuo shenme? *'What did you say?'*	ni shuo shenme? *'What did you say?'*	ni shuo shenme? *'What did you say?'*
	A	wuma mo maomi. *Mom Wu patted the kitty.*	laomo na maoni. *The model worker took the woolen.*	limei lou laoma. *Li Mei hugged Lao Ma.*	luli ma laile. *Lu Li scolded Lai Le.*
Initial focus	Q	shui mo maomi? *'Who patted the kitty?'*	shui na maoni? *'Who took the woolen?'*	shui lou laoma? *'Who hugged Lao Ma?'*	shui ma laile? *'Who scolded Lai Le?'*
	A	wuma mo maomi.	laomo na maoni.	limei lou laoma.	luli ma laile.
Medial focus	Q	wuma dui maomi zuo shenme? *'What did Mom Wu do to the kitty?'*	laomo dui maoni zuo shenme? *'What did the model worker do to the woolen?'*	limei dui laoma zuo shenme? *'What did Li Mei do to Lao Ma?'*	luli dui laile zuo shenme ? *'What did Lu Li do to Lai Le?'*
	A	wuma mo maomi.	laomo na maoni.	limei lou laoma.	luli ma laile.
Final focus	Q	wuma mo shenme? *'What did Mom Wu pat?'*	laomo na shenme? *'What did the model worker take?'*	limei lou shui? *'Who did Li Mei hug?'*	luli ma shui ? *'Who did Lu Li scold?'*
	A	wuma mo maomi.	laomo na maoni.	limei lou laoma.	luli ma laile.

TABLE II. TARGET SENTENCES IN STANDARD, NANJING, CHANGCHUN MANDARINS

	Word 1	Word 2	Word 3
Character	邬妈	摸	猫咪
Gloss	'mom Wu'	'pat'	'kitty'
Standard	wu55 ma55	mo55	mao55 mi55
Nanjing	wu31 ma31	mo31	mao31 mi31
Changchun	wu44 ma44	mo44	mao44 mi44
Character	劳模	拿	毛呢
Gloss	'model worker'	'take'	'woolen'
Standard	lao35 mo35	na35	mao35 ni35
Nanjing	lao13 mo13	na13	mao13 ni13
Changchun	lao24 mo24	na24	mao24 ni24
Character	李美	搂	老马
Gloss	'Li Mei'	'hug'	'Lao Ma'
Standard	li214 mei214	lou214	lao214 ma214
Nanjing	li212 mei212	lou212	lao212 ma212
Changchun	li214 mei214	lou214	lao214 ma214
Character	陆丽	骂	赖乐
Gloss	'Lu Li'	'scold'	'Lai Le'
Standard	lu51 li51	ma51	lai51 le51
Nanjing	lu44 li44	ma44	lai44 le44
Changchun	lu53 li53	ma53	lai53 le53

C. *Procesures*

Recording took place in a soundproof booth at NJUST. The two experimenters were conducted by a native speaker of CC and NJ Mandarins respectively. Prompt questions were asked in a random order with five repetitions. The written answers were provided in Chinese characters. Participants were requested to respond to the questions with appropriate foci. A total of 80 target sentences were produced by each participant (4 tones × 4 focus locations × 5 repetitions). A Marantz professional solid state recorder PMD661 and a Shure professional unidirectional head-worn dynamic microphone SM10A-CN were used for recording. The stimuli were recorded with a sampling rate of 44,100HZ and saved in an SD card. The production experiment was self-paced and lasted approximately 30 minutes for each participant.

D. *Analysis*

Data were analyzed by Praat version 5.3.65 [8] and a Praat scritpt—ProsodyPro version 5.5.2 [23]. The acoustic parameters in the current paper include mean duration, mean F0, mean intensity and time-normalized F0 at ten even-interval points in each syllable. The ten even-interval F0 values were extracted to track the time-normalized F0 trajectory for a direct observation of pitch change as a function of focus.

To examine the prosodic change of focus, the differential of duration, F0 and intensity values between the initial, medial, final foci and the neutral focus was calculated by subtracting the mean value of each component in the neutral-focus sentence from the corresponding mean value in the focused sentences (initial, medial, and final respectively). Note that the mean F0 differential was converted from Hertz to semitones [st = 12 log2(F0), where reference level is 1 Hz] because pitch in speech operates on a logarithmic scale just as in music [5, 6].

Therefore, pre-focus change was calculated on, for instance, the mean value of the two syllables of "wuma" in the medial-focus sentences and the three syllables of "wuma mo" in the final-focus sentences minus their counterparts in the neutral-focus sentences. The in-focus change was computed on the prosodic values of the syllables of the focused words minus its neutral-focus counterparts. Post-focus change was calculated on the mean value of the three syllables of "mo maomi" in the initial-focus sentences and the two syllables of "maomi" in the medial-focus sentences minus that of the neutral-focus sentences.

III. RESULTS

A. *Time-normalized F0 Contours*

Figures 1-8 display the time-normalized F0 contours with three focus locations (initial, medial, final) vs. their neutral-focus counterparts by tone and dialect. Each curve represents an average of each target sentence of five repetitions produced by the ten speakers of each group. The focused sentences are represented by dash curves, while the neutral-focus sentences by solid curves. Syllable boundaries are signified by vertical lines.

Figures 1-4 display the time-normalized F0 contours in four tones in CC Mandarin produced by CC speakers. It reveals that they produced clear in-focus expansion in F0 in the three focus locations for four tones and post-focus compression in initial and medial focus locations. The pre-focus changes are not as clear as in-focus and post-focus changes.

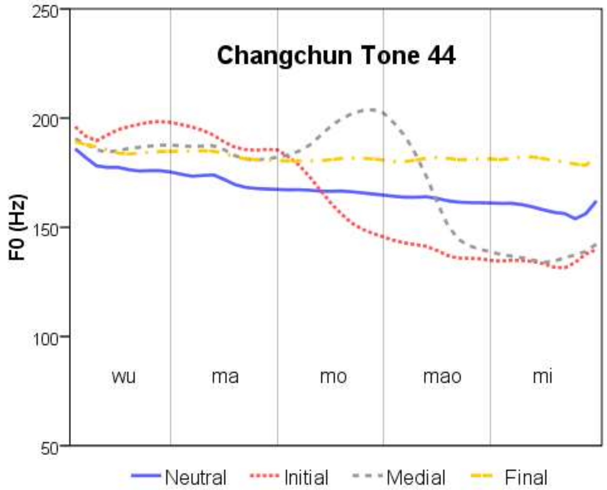

Figure 1. Time-normalized F0 contours (Hz) by Changchun speakers in Tone 44

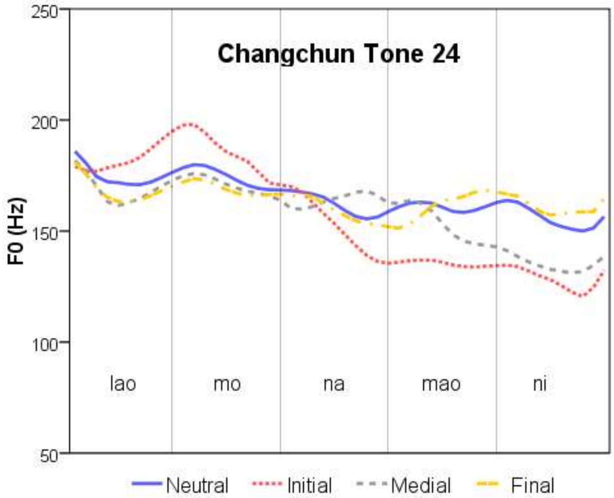

Figure 2. Time-normalized F0 contours (Hz) by Changchun speakers in Tone 24

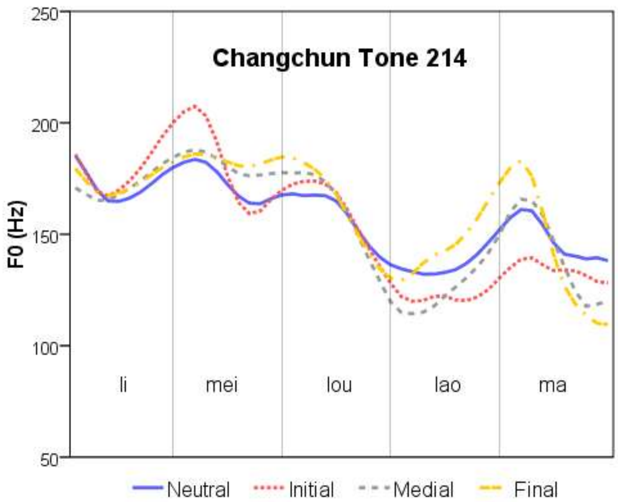

Figure 3. Time-normalized F0 contours (Hz) by Changchun speakers in Tone 214

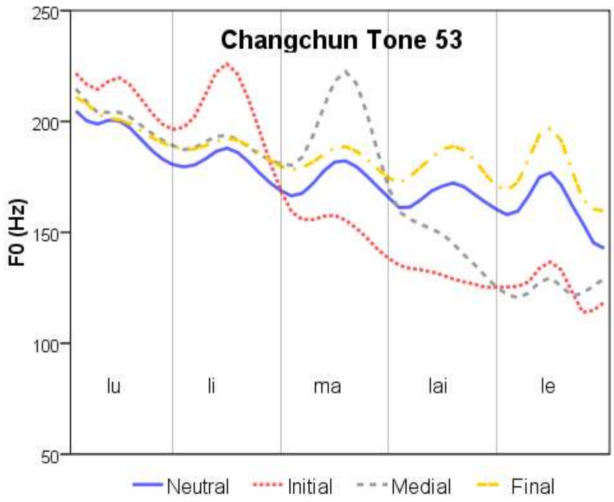

Figure 4. Time-normalized F0 contours (Hz) by Changchun speakers in Tone 53

Likewise, the in-focus expansion was quite noticeable in NJ Mandarin as shown in Figures 5-8. The F0 contours also show noticeable post-focus compression in the initial and medial focus locations with the four tones whereas unnoticeable change in pre-focus constituents.

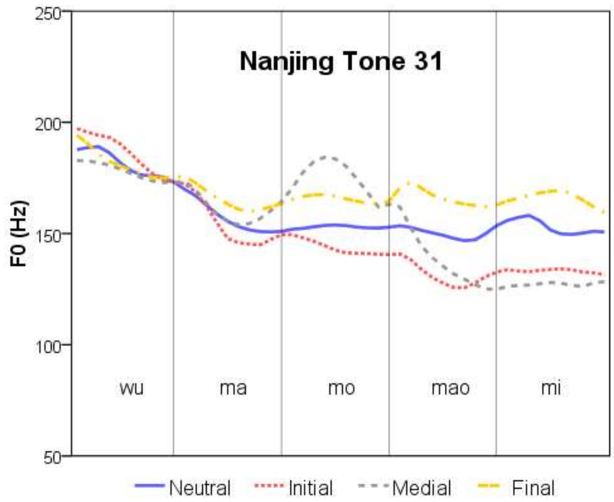

Figure 5. Time-normalized F0 contours (Hz) by Nanjing speakers in Tone 31

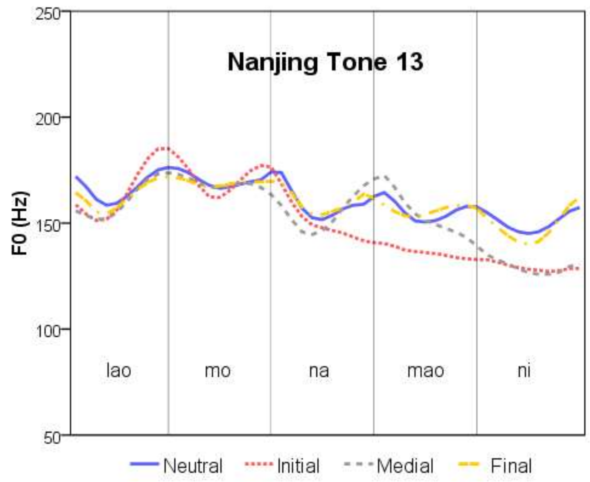

Figure 6. Time-normalized F0 contours (Hz) by Nanjing speakers in Tone 13

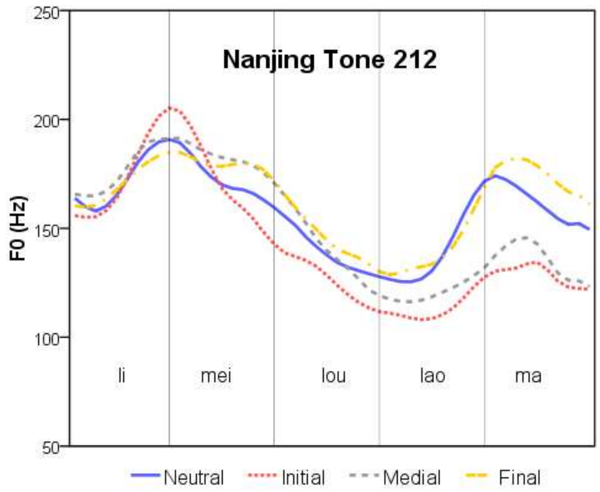

Figure 7. Time-normalized F0 contours (Hz) by Nanjing speakers in Tone 212

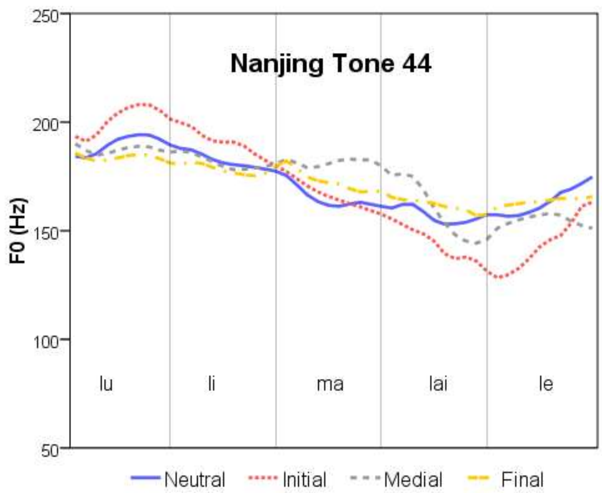

Figure 8. Time-normalized F0 contours (Hz) by Nanjing speakers in Tone 44

B. Focus Change in Mean F0, Intensity, and Duration

The time-normalized F0 contours illustrate consistent patterns of prosodic change as a function of focus regardless of tone. Therefore, focus changes were calculated collapsing tone types in the current paper. We applied repeated measures ANOVAs to analyze the mean F0, intensity, and duration change with a within-subjects factor—focus condition (three levels: pre-focus, in-focus, and post-focus) and a between-subject factor—speaker group (two levels: Changchun and Nanjing).

The ANOVA results of mean F0 change indicate no interaction between focus condition and speaker group and the main effect of the speaker group show no significance. The main effect of focus condition ($F(2, 36) = 84.181$, $p < 0.001$) was highly significant. The results demonstrate that there was no significant difference between CC and NJ groups. Further comparisons in paired-samples *t*-tests show significant differences between pre-focus change and post-focus change ($t(9) = 5.097$, $p = 0.001$) and between in-focus change and post-focus change ($t(9) = 7.238$, $p < 0.001$) in NJ Mandarin. Similar comparisons also show significant differences between pre-focus change and in-focus change ($t(9) = -4.628$, $p = 0.001$), between pre-focus change and post-focus change ($t(9) = 7.660$, $p < 0.001$), and between in-focus change and post-focus change ($t(9) = 8.635$, $p < 0.001$) in CC Mandarin.

To compare the degree of using F0 change to code focus at sentential level between CC and NJ speakers, the magnitude on F0 differentials was computed. It respectively subtracted pre-focus change and post-focus change from in-focus change and was examined by independent-samples *t*-tests. Results show no significant difference between CC and NJ speakers in F0 change magnitude between in-focus and pre-focus ($t(18) = -0.387$, $p = 0.703$) and between in-focus and post-focus ($t(18) = -0.612$, $p = 0.548$).

Figure 9 displays mean F0 change with standard error bars of the two speaker groups. CC and NJ speakers expanded F0 in focused components and compressed F0 in post-focus components (thus produced PFC) whereas no consistent change in pre-focus condition.

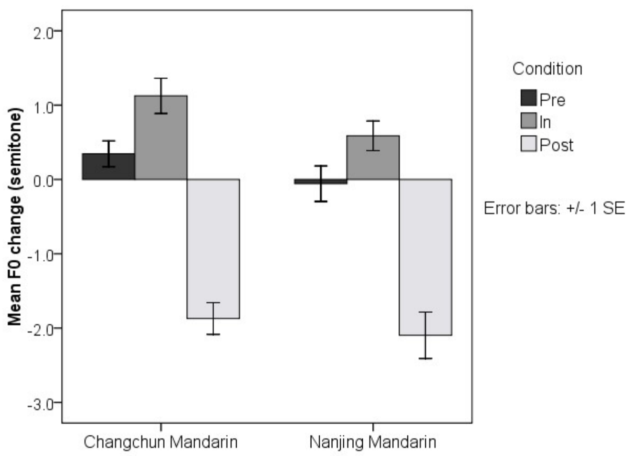

Figure 9. F0 change (semitone) by focus condition and speaker group

The results of repeated measures ANOVA on mean intensity change indicate no interaction between focus condition and speaker group, no main effect of speaker group and only main effect of focus condition ($F(2, 36) = 119.151$, $p < 0.001$). Further comparisons in paired-samples *t*-tests show significant differences between pre-focus change and in-focus change ($t(9) = -3.785$, $p = 0.004$), between pre-focus change and post-focus change ($t(9) = 7.259$, $p < 0.001$) and between in-focus change and post-focus change ($t(9) = 11.134$, $p < 0.001$) in NJ Mandarin. And significant differences between pre-focus change and in-focus change ($t(9) = -4.476$, $p = 0.002$), between pre-focus change and post-focus change ($t(9) = 5.582$, $p < 0.001$) and between in-focus change and post-focus change ($t(9) = 9.452$, $p < 0.001$) were found in CC Mandarin.

Independent-samples *t*-tests show no significant difference between CC and NJ speakers in intensity change magnitude between in-focus and pre-focus ($t(18) = -0.419$, $p = 0.680$) and between in-focus and post-focus ($t(18) = 1.174$, $p = 0.256$). Figure 10 reconfirms these results that CC and NJ speakers produced clear in-focus expansion and post-compression of intensity and there is no difference between CC and NJ.

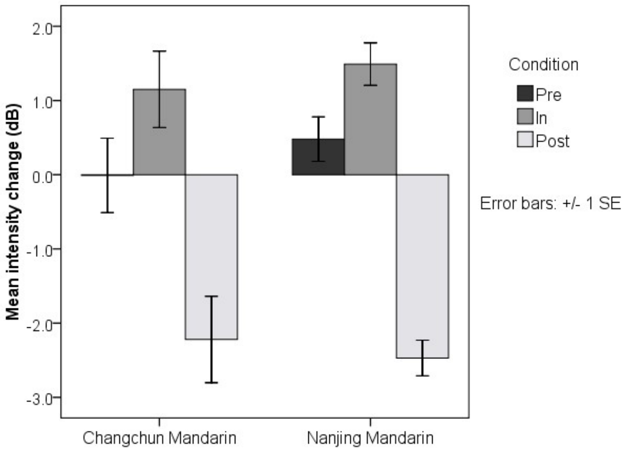

Figure 10. Intensity change (dB) by focus condition and speaker group

The results of repeated measures ANOVA on mean duration change demonstrate again only main effect of focus condition was significant ($F(1.40, 25.27) = 31.488$, $p < 0.001$) but no interaction between focus condition and speaker group and no main effect of group. Paired-samples *t*-tests show significance between pre-focus change and in-focus change ($t(9) = -3.298$, $p = 0.009$) and between in-focus change and post-focus change ($t(9) = 2.868$, $p = 0.019$) in NJ Mandarin, and significance between pre-focus change and in-focus change ($t(9) = -6.072$, $p < 0.001$) and between in-focus change and post-focus change ($t(9) = 6.203$, $p < 0.001$) in CC Mandarin.

Independent-samples *t*-tests also showed no significant difference between CC and NJ speakers in duration change magnitude between in-focus and pre-focus ($t(18) = -1.269$, $p = 0.221$) and between in-focus and post-focus ($t(18) = -1.037$, $p = 0.314$). As shown in Figure 11, CC and NJ speakers produced clear in-focus expansion in duration while there was no clear patterns of duration change in pre-focus and post-focus components.

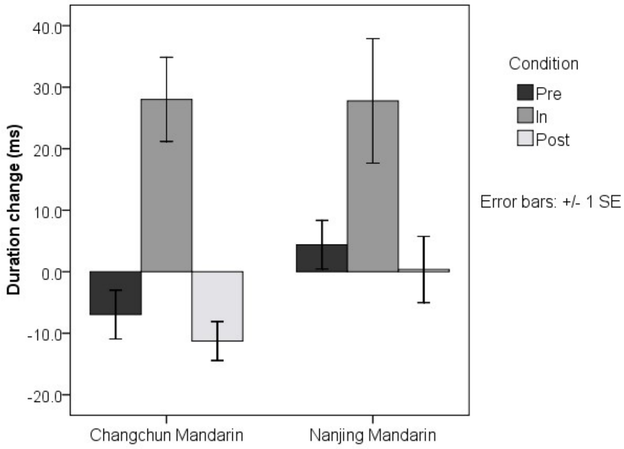

Figure 11. Duration change (ms) by focus condition and speaker group

IV. Discussion

The results of the speech production experiment have answered the three research questions. For the first and second questions, both CC and NJ Mandarin speakers produced in-focus expansion of F0, intensity and duration and post-focus compression of F0 and intensity. However, they did not produce substantial pre-focus change in either CC Mandarin or NJ Mandarin. This finding is in line with that of [21] that the prosodic change in pre-focus components normally remains intact.

According to the results of ANOVAs and *t*-tests, there was no significant difference in the prosodic change of mean F0, intensity and duration as a function of focus between CC Mandarin and NJ Mandarin, which provides negative answers to the third research question. In addition, no difference of the magnitude of F0, intensity and duration between in-focus and pre-focus changes and between in-focus and post-focus changes in CC and NJ Mandarin productions suggests that both speaker groups use similar degree of acoustic variations to code focus in their Mandarin dialects though they differ in tonal inventory. This further suggests that prosodic variation is a means of coding focus at sentential level regardless of tones at lexical level.

V. Conclusions

This paper examines the prosodic realization of focus in Mandarin dialects by Changchun and Nanjing Mandarin speakers. Two main findings are discovered. First, both CC and NJ speakers produced in-focus expansion and post-focus compression (PFC) in the two Mandarin dialects. Second, CC and NJ speakers use similar prosodic pattern to code focus in these two dialects. We conclude that Changchun Mandarin and Nanjing Mandarin are PFC language.

PFC has been found not easy to transfer from one language to another [15, 16, 17], even highly proficient Chinese learners of English were not able to achieve native-like PFC in English and vice versa, future work may investigate the English production of bidialectal speakers of standard Mandarin and Changchun/Nanjing Mandarin for insight into whether one more PFC dialect in learners' L1 facilitates acquiring PFC in their L2.

ACKNOWLEDGMENT

This work was supported by the National Science Foundation of China (61573187) and the NJUST Key Bachelor's Thesis Grant.

REFERENCES

[1] B. Wang and Y. Xu, "Differential prosodic encoding of topic and focus in sentence-initial position in Mandarin Chinese," *Journal of Phonetics*, vol. 39, pp. 595-611, 2011.

[2] C. Féry and F. Kügler, "Pitch accent scaling on given, new, and focused constituents in German," *Journal of Phonetics*, vol. 36, pp. 0-703, 2008.

[3] C. T. Best and M. D. Tyler, "Nonnative and second-language speech perception: Commonalities and complementarities," Ed. *Language experience in second language speech learning: In honor of James Emil Flege.* Amsterdam: John Benjamins Publishing Company, pp. 13-34, 2007.

[4] C. T. Best, "A Direct Realist View of Cross-Language Speech Perception," Ed. *Speech perception and linguistic experience: Issues in cross-language research.* Timonium, MD: York Press, pp. 171-204, 1995.

[5] F. Nolan, "Intonational equivalence: an experimental evaluation of pitch scales," In *Proceedings of the 15th international congress of phonetic sciences*, Barcelona, Spain, vol. 771, pp. 774, 2003.

[6] H. Fujisaki, "Prosody, information, and modeling-with emphasis on tonal features of speech," *Workshop on Spoken Language Processing*, pp. 5-14, 2003.

[7] J. E. Flege, "Second-language speech learning: Theory, findings, and problems," Ed., *Speech perception and linguistic experience: Issues in cross-language research.* Timonium, MD: York Press, pp. 233-277, 1995.

[8] P. Boersma and D. Weenink, "Praat: doing phonetics by computer. Version 5.3.65. Online: http://www.fon.hum.uva.nl/praat/," 2014.

[9] R. Lado, "Linguistics across cultures: applied linguistics for Language teachers," University of Michigan Press, 1957.

[10] R. Wardhaugh, "The contrastive analysis hypothesis," *TESOL quarterly*, pp. 123-130, 1970.

[11] S. W. Chen, B. Wang and Y. Xu, "Closely related languages, different ways of realizing focus," In *Proceedings of Interspeech*, Brighton, UK, pp. 1007-1010, 2009.

[12] S. Zerbian, E. O. Aboh, K. Hartmann and M. Zimmermann, "Investigating prosodic focus marking in Northern Sotho," *Focus strategies in African languages: The interaction of focus and grammar in Niger-Congo and Afro-Asiatic*, pp. 55-79, 2007.

[13] T. V. Perneger, "What's wrong with Bonferroni adjustments," *Bmj, 316*(7139), pp. 1236-1238, 1998.

[14] W. L. Wu and Y. Xu, "Prosodic focus in Hong Kong Cantonese without post-focus compression," *Speech Prosody 2010-Fifth International Conference*, Chicago, USA, pp.148-151, 2010.

[15] Y. Chen, "Post-focus compression in English by Mandarin learners," *Proceedings of the 18th International Congress of Phonetic Sciences*, pp. 1-5, 2015.

[16] Y. Chen, "Prosodic comparisons of two types of realization of focus in Mandarin," *Speech Prosody*, pp. 917-921, 2018.

[17] Y. Chen, S. Guion-Anderson and Y. Xu, "Post-focus compression in second language Mandarin," *Speech Prosody*, pp. 410-413, 2012.

[18] Y. Chen, Y. Xu and S. Guion-Anderson, "Prosodic realization of focus in a bilingual production of Southern Min and Mandarin," *Phonetica*, vol. 71, pp. 249-270, 2014.

[19] Y. R. Chao, "A system of tone letters." *Le maître phonétique*, pp. 24-27, 1930

[20] Y. Xu and C. X. Xu, "Phonetic realization of focus in English declarative intonation," *Journal of Phonetics*, vol. 33, pp. 159-197, 2005.

[21] Y. Xu, "Effects of tone and focus on the formation and alignment of F0 contours," *Journal of Phonetics*, vol. 27, pp. 55-105, 1999.

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

[23] Y. Xu, "ProsodyPro.praat. Version 5.5.2. Online: https://www.phon.ucl.ac.uk/home/yi/ProsodyPro/," 2014.

[24] Y. Xu, S. W. Chen and B. Wang, "Prosodic focus with and without post-focus compression (PFC): A typological divide within the same language family?" *The Linguistic Review*, pp. 131-147, 2012.

[25] F. L. Wang and S. Q. Li, "A Survey of the Changchun Dialect." Jilin Literature and History Publishing house, 1995.

[26] Nanjing Local Chronicle Compilation Committe and Dialect Chronicle Compilation Committee, "A Survey of the Nanjing Dialect." Nanjing Press, 1993.