Chinese EFL Learners' Perception of English Prosodic Focus

*Xinya Zhang, Ying Chen**

MIIT Key Lab for Language Information Processing and Applications
School of Foreign Studies, Nanjing University of Science and Technology, China
zhang_xinya@njust.edu.cn, ychen@njust.edu.cn*

Abstract

Focus in a sentence can be realized prosodically in speech communication. It has been found not easy for L2 learners to acquire. The present study examines Chinese learners' perception of English prosodic focus, specifically the effects of learners' English proficiency, intonation type, sentence length, and focus location on the perceptual accuracy of English prosodic focus by Chinese EFL learners. Results of two trials in the perception experiment reveal that focus location, intonation type, and English proficiency significantly impacted Chinese learners' perceptual accuracy of both single focus and dual focus in English. Focus in statements was perceived more accurately than that in questions for both single focus and dual focus. Focus located on sentence-final words in questions was perceived more accurately than that on non-final words in questions. Learners' English proficiency positively correlated to the accuracy of focus perception, especially for dual focus.

Index Terms: prosodic focus, perception, EFL, single focus, dual focus, post-focus compression

1. Introduction

Focus refers to a certain part of an utterance that contains information more salient than others. Focus can be used to make a contrast/correction in a statement, or to provide new information for a WH-question [1, 2]. Prosody is one means of focus marking [3–7].

Compared to its unfocused counterparts, prosodic focus is in general encoded with an in-focus expansion in fundamental frequency (F0), duration, and intensity, a post-focus compression in F0 and intensity, and no consistent prosodic change in pre-focus words in many languages [5, 6, 8]. Post-focus compression (PFC) was attested in American English [3, 6], British English [9], Finnish [10], Korean [11], Dutch [12], Persian [13], Beijing Mandarin [5, 7], Nanjing Mandarin and Changchun Mandarin [14], while PFC was absent in Hong Kong Cantonese [9, 15], Taiwan Southern Min [7], and Quanzhou Southern Min [16]. Hence, PFC is a non-universal cue of focus marking [17]. In-focus expansion was considered as a predominant cue to focus perception [18], and focus location was found to be identified more accurately if PFC was present [7, 8, 10, 12, 17, 18].

Dual focus is defined as the presence of two foci in a mono-clausal sentence [19, 20]. For example, in the answer to the question "Who ate what in the restaurant?" "JOHN ate CHILI in the restaurant.", "John" and "chili" are focused [1]. Prosodic realization was found different between the early and late focus for dual focus [12, 19–23]. In English dual-focus statements, the early focus lacked a lowered post-focal F0, indicating an anticipatory influence of the late focus—the F0 increase of the late focus was facilitated by minimizing the F0 drop in post-

early-focus items [21]. For the prosodic correlates of dual focus (initial+final focus) in Mandarin, both foci increased F0 and word duration to almost the same degree as its initial and final focus counterpart. However, the word after the early focus was not different from its neutral counterpart in either F0 or duration [23].

Intonation type is another factor affecting prosodic realization of focus [2, 24]. Pertaining to English focus in statements (with low/falling pitch) and yes/no questions (with high/rising pitch), questions showed overall higher F0 values than corresponding statements and a particularly high value at the end of the sentence was found in final-focus questions [25]. In [24], native speakers of Mandarin were able to perceive statement/question contrast and the presence and location of focus simultaneously in Mandarin, whereas low identification accuracy was observed for neutral focus (i.e. non-narrow focus) in questions and for statements with final focus.

It has been found difficult for L2 learners to perceive and produce prosodic focus in a native-like fashion. There was loss of PFC through language contact, e.g., in Taiwan Mandarin produced by Taiwan Southern Min-Mandarin bilinguals [7] and in English produced by English-Cantonese bilinguals [9]. Wang and colleagues [26] investigated the perception of Mandarin prosodic focus by native speakers of Mandarin—a PFC language, and by native speakers of Tsat—a non-PFC language. The results indicated that the Tsat speakers, who were L2 speakers of Mandarin, perceived focus with much lower accuracy than L1 Mandarin speakers [26]. As one of the best-known models concerning L2 speech acquisition, the Speech Learning Model (SLM) [27] and its revised version SLM-r [28] suggest that although L2 learners may not be able to perfectly produce L2 sounds as monolingual native speakers do, individuals differ in terms of how accurately they perceive and produce L2 sounds. Pertaining to prosodic focus, Chen and associates [16, 29, 30] noted that PFC was learnable given sufficient language experience, e.g., early age of learning (AOL) an L2, an extensive amount of L2 use, greater length of residence in a predominantly L2-speaking environment, and high quantity and quality of L2 input [31–35].

Although both Mandarin and English are PFC languages [3, 5–7], Mandarin learners of English, even highly proficient, were unable to produce native-like focus in English [29, 36]. However, their perception of English prosodic focus remains hitherto unknown. Three research questions are therefore explored in the present study:

1) Do Chinese EFL learners perceive English single and dual focus accurately?

2) Is the perceptual accuracy of focus in questions different from that in statements in English by Chinese learners?

3) Does the English proficiency of Chinese EFL learners impact their perception of English prosodic focus?

2. Methods

2.1. Participants

Eighty undergraduates (mean age 21.2 yrs. with SD 1.3) were recruited in China as listeners. All the participants are native speakers of Mandarin from northern China, learning English as a foreign language (EFL), and with no self-reported hearing or speech disorders. They were non-English majors and had taken the College English Test Band-6 (CET-6), a national English-as-a-foreign-language test in China, by the time of the experiment. The participants were divided into two groups—40 learners with high English proficiency who had passed CET-6, and 40 learners with low proficiency who had not passed. Each group had 20 females and 20 males.

2.2. Stimuli

Six sentences with three different lengths served as stimuli—two short sentences with 5 syllables, two medium sentences with 8 syllables, and two long sentences with 12 syllables. The initial, medial, and/or final content words were designed to be focused (see Table 1).

Table 1: *Target sentence list.*

	Initial		Medial		Final
Short	**Ray**	may	**know**	my	**mom**
	Lee	may	**wear**	my	**ring**
Medium	**Molly**	will	**marry**	my	**lawyer**
	Laura	will	**ruin**	my	**morning**
Long	**Nina**	wanna mail	**Larry**	a yellow	**lemon**
	Mary	wanna loan	**Lily**	a normal	**mirror**

Each sentence had 14 versions differing in focus location (neutral, initial, medial, final, initial+medial, initial+final, medial+final) and intonation type (statement, question). Stimuli were recorded by two native speakers of American English (one female, age 30 yrs. and one male, age 34 yrs.) with a 44,100 Hz sampling rate. The talkers were requested to read aloud the target sentences presented in a random order as naturally as possible and at a normal speaking rate and emphasize the font-highlighted words that were focused in the sentences. For neutral focus, they were told not to emphasize any word. The talkers were instructed to say each sentence as either a statement or a question depending on whether it ended with a period or a question mark. Each talker was paid $100 for recording 84 English sentences (6 sentences × 7 focus locations × 2 intonation types). Another two native speakers of American English (one female, age 35 yrs. and one male, age 26 yrs.) were paid $40 each to listen to the sentences and make sure that the stimuli were produced naturally and prosodically with focus.

2.3. Procedure

The stimuli were presented to listeners also in a randomized order at a comfortable and fixed loudness level (see [10] for intensity effects on focus perception) via E-Prime 3.0. The participants were requested to identify the presence and locations of focus, which was instructed as emphasis before the experiment [8, 10, 12, 24, 37].

The experiment consisted of two trials for single and dual focus respectively. In both trials, four response categories, i.e., specific words in initial/medial/final positions and "none" were displayed on the computer screen, and listeners chose the one/two that matched his/her impression by pressing keys on the keyboard. For single focus, the keys "1", "2", "3", and "4" represented initial, medial, final, and neutral focus respectively. For dual focus, double pressing of "1" + "2", "1" + "3", "2" + "3", and "0" + "0" represented initial+medial, initial+final, medial+final, and neutral focus respectively. The question "which word/two words was/were emphasized" in English was presented visually to the participants. Each stimulus sentence was played once and there was no time limitation for the participants to make a judgment. This perception experiment lasted about 30–40 minutes. Each participant was paid ¥40. In total, 192 responses (2 trials × 6 sentences × 2 intonation types × 4 focus locations × 2 talkers) were collected from each listener.

2.4. Data collection and analysis

Response and accuracy ("correct" coded as "1", "incorrect" as "0") were collected via E-Prime 3.0. R (version 4.2.2) was used for statistical analysis. Mixed-effects logistic regression model was implemented by the *glmer* function in the "lme4" package for the results of single and dual focus separately. The dependent variable was listeners' response ("0" for inaccurate, "1" for accurate). Learner group (2 levels: high English proficiency, low English proficiency), intonation type (2 levels: statement, question), sentence length (3 levels: short, medium, long), and focus location (4 levels: neutral, initial, medial, final for single focus; neutral, initial+medial, initial+final, medial+final for dual focus) were set as fixed factors. Participant, talker, and sentence were set as random factors.

To better illustrate the results of two mixed-effects logistic regression models, the "afex" package was used to calculate *p*-values for all fixed effects using likelihood-ratio tests. When a significant effect was observed, the *emmeans* function in the "emmeans" package was further used to conduct Tukey-HSD pairwise comparisons. The "ggplot2" package was employed to create data graphics [38]. All statistical effects are reported at a significance level of 0.05.

3. Results

3.1. Single focus

As listed in Table 2, results demonstrate significant main effects of learner group, intonation type and focus location on Chinese learners' perceptual accuracy of single focus in English. Sentence length exerted no significant main effect on the perception of single focus [$\chi^2(2) = 0.56$, $p = 0.757$]. Five two-way interactions and a three-way interaction demonstrated statistical significance.

In Figures 1 and 2, the percent accuracy of learners' perception of single focus in English statements and questions was plotted respectively by learner group and focus location, collapsing sentence length. Highly-proficient learners perceived focus significantly better than low-proficient learners under all the focus locations in both intonation types ($ps < 0.05$) except for initial focus in questions (High vs. Low proficiency: $\beta = -0.068$, $SE = 0.202$, $z = -0.335$, $p = 0.737$) and final focus in questions (High vs. Low proficiency: $\beta = 0.348$, $SE = 0.238$, $z = 1.461$, $p = 0.144$). The confusion matrix of single focus perception is illustrated in Table 3.

Table 2: *Mixed-effects logistic regression results of Chinese EFL learners' perception of single focus.*

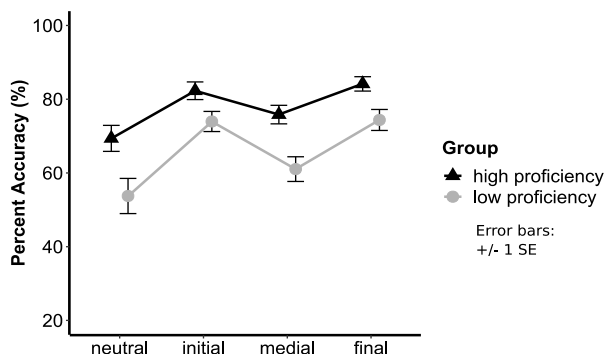| Effect	*Df*	χ^2	*p*
Group	1	10.95	<0.001
Intonation	1	114.99	<0.001
Focus	3	372.75	<0.001
Group × Intonation	1	8.24	0.004
Length × Intonation	2	21.11	<0.001
Group × Focus	3	10.03	0.018
Length × Focus	6	79.25	<0.001
Intonation × Focus	3	81.18	<0.001
Length × Intonation × Focus	6	39.58	<0.001

Figure 1: *Percent accuracy of Mandarin learners' perception of single focus in English statements.*
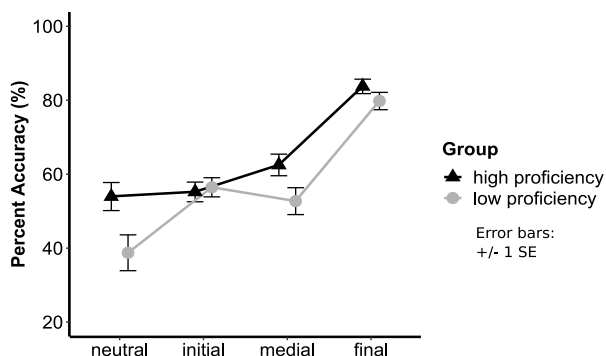

Figure 2: *Percent accuracy of Mandarin learners' perception of single focus in English questions.*

Table 3: *Confusion matrix of classification percentage (%) for single focus (S = Statement, Q = Question).*

Listeners' proficiency	Spoken		Heard			
			Neutral	Initial	Medial	Final
High	S	Neutral	**69.2**	20.6	4.2	6.0
		Initial	13.1	**82.3**	1.3	3.3
		Medial	12.5	5.4	**75.8**	6.3
		Final	9.4	5.2	1.3	**84.2**
High	Q	Neutral	**54.0**	10.6	8.5	26.9
		Initial	27.1	**55.2**	2.5	15.2
		Medial	19.6	5.2	**62.5**	12.7
		Final	11.7	2.3	2.3	**83.8**
Low	S	Neutral	**53.8**	25.6	7.3	13.3
		Initial	11.5	**74.0**	5.0	9.6
		Medial	10.8	14.4	**61.0**	13.8
		Final	8.3	12.7	4.6	**74.4**

Table 3 (continued).

Low	Q	Neutral	**38.8**	16.9	9.8	34.6
		Initial	14.4	**56.5**	6.3	22.9
		Medial	16.7	11.3	**51.7**	20.4
		Final	9.4	4.6	6.3	**79.8**

3.2. Dual focus

Table 4 illustrates significant main effects of group, intonation, and focus on Mandarin learners' perceptual accuracy of English dual focus. Sentence length exerted no significant main effect on the perceptual accuracy of English dual focus [$\chi^2(2) = 2.66$, $p = 0.265$]. There were five two-way interactions and a three-way interaction that showed statistical significance.

Table 4: *Mixed-effects logistic regression results of Chinese EFL learners' perception of dual focus.*

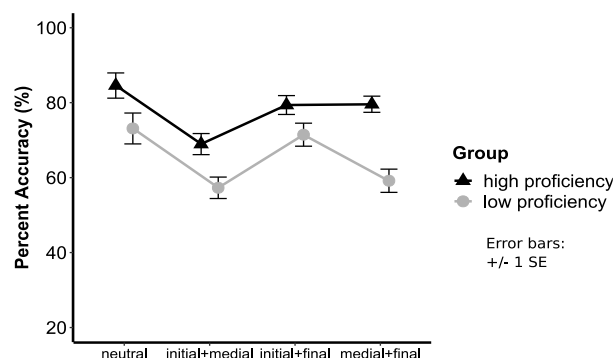| Effect	*Df*	χ^2	*p*
Group	1	17.09	<0.001
Intonation	1	74.82	<0.001
Focus	3	136.65	<0.001
Group × Intonation	1	6.22	0.013
Length × Intonation	2	8.66	0.013
Group × Focus	3	20.84	<0.001
Length × Focus	6	57.57	<0.001
Intonation × Focus	3	68.50	<0.001
Length × Intonation × Focus	6	58.93	<0.001

Figure 3: *Percent accuracy of Mandarin learners' perception of dual focus in English statements.*
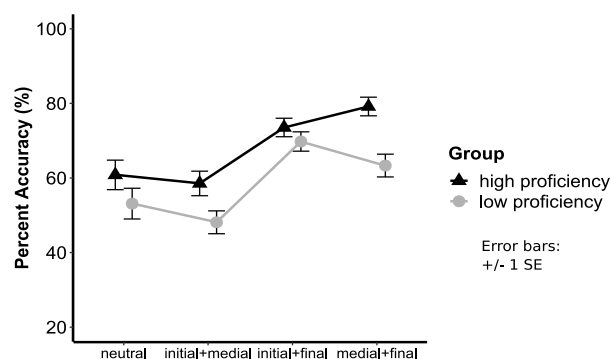

Figure 4: *Percent accuracy of Mandarin learners' perception of dual focus in English questions.*

In Figures 3 and 4, the percent accuracy of Chinese learners' perception of dual focus in English statements and questions was plotted respectively by learner group and focus location,

collapsing sentence length. Chinese learners with high English proficiency perceived focus better than those with low English proficiency under all focus locations in both intonation types ($ps < 0.05$) except for neutral focus in questions ($\beta = 0.369$, $SE = 0.196$, $z = 1.884$, $p = 0.060$) and initial-final focus in questions ($\beta = 0.217$, $SE = 0.205$, $z = 1.056$, $p = 0.291$). The confusion matrix of dual focus perception is shown in Table 5.

Table 5: *Confusion matrix of classification percentage (%) for dual focus (S = Statement, Q = Question, Neutral = N, Initial+medial = I+M, Initial+final = I+F, Medial+final = M+F).*

Listeners' proficiency	Spoken		Heard			
			N	I+M	I+F	M+F
High	S	N	**84.6**	5.2	6.7	3.5
		I+M	7.1	**69.0**	10.0	14.0
		I+F	6.0	4.0	**79.2**	10.8
		M+F	8.1	4.6	7.7	**79.6**
High	Q	N	**60.8**	9.0	22.1	8.1
		I+M	14.4	**58.5**	13.3	13.8
		I+F	11.7	4.2	**73.5**	10.6
		M+F	6.9	6.7	7.1	**79.4**
Low	S	N	**73.1**	7.3	14.2	5.4
		I+M	9.2	**57.3**	21.7	11.9
		I+F	9.0	7.5	**71.3**	12.3
		M+F	11.3	11.5	18.1	**59.2**
Low	Q	N	**53.1**	10.4	25.6	10.8
		I+M	15.0	**48.1**	18.5	18.3
		I+F	11.0	8.8	**69.8**	10.4
		M+F	12.3	9.0	15.4	**63.3**

4. Discussion

In line with the previous research [26], the results in the current study respond to the first research question that Chinese learners may have difficulty perceiving English prosodic focus accurately. Chinese learners' perceptual accuracy, according to the statistical analysis, is significantly impacted by focus location for both single focus and dual focus.

For single focus, final focus was more accurately perceived than neutral, initial and medial focus in questions with all sentence lengths ($ps < 0.001$). For dual focus in all sentence lengths, initial+final focus received higher accuracy than initial+medial focus in questions ($ps < 0.01$) and medial+final focus received also higher accuracy than initial+medial focus in questions ($ps < 0.001$). Based on the Parallel Encoding and Target Approximation (PENTA) model [39–42], focus and intonation are encoded in parallel and both yes/no question and final focus are realized with a pitch raising in the sentence-final position in English. Hence, the low perceptual accuracy of focus located on non-final words might be explained by the confusions caused by the similar encoding characteristics between final focus and question. When the F0 of the final word in a question was raised based on the interrogative meaning and somewhat resembling the F0 pattern caused by final focus [22, 24, 25, 37], the listeners might be confused and attributed the F0 raising to focus and thus considered it sentential-final focus, which includes final focus, initial+final focus, or medial+final focus.

The results of the present study also provide a positive answer to the second research question that Mandarin learners did perceive focus in questions with different accuracy from that in statements. Focus in statements was better perceived

than that in questions in all focus locations ($ps < 0.05$) except for final focus (Question vs. Statement: $\beta = 0.220$, $SE = 0.127$, $z = 1.725$, $p = 0.085$) and medial+final focus (Question vs. Statement: $\beta = 0.040$, $SE = 0.114$, $z = 0.348$, $p = 0.728$). This lower perceptual accuracy of focus in questions, as mentioned above, might be attributed to the low accuracy of focus located on non-final words in questions due to confusions caused by the similar encoding schemes in shaping F0 contours between final focus and question.

Another question of particular interest concerned the role of English proficiency in learners' perceptual accuracy of English prosodic focus. As reported in the Results section, Chinese learners with high English proficiency overall perceived focus location with significantly higher accuracy than those with low English proficiency, for dual focus in particular, for which learners with high proficiency scored numerically higher than learners with low proficiency in all focus locations in both intonation types. Learners with higher English proficiency might have been more interested and with stronger motivation in learning English. Therefore, they may have intentionally received more input from native English speakers via mass media (films, television, radio, the Internet, etc.) and thus with more language experience than learners with low English proficiency, which might result in their better perception of English focus. Echoing previous findings [16, 29, 30] that more native-like production of prosodic focus and PFC could be learned through sufficient language experience in L2/foreign language, the current study demonstrates that with L2/foreign language proficiency improving, learners are able to perceive prosodic focus more accurately as well and thus produce it more native-like. These findings are consistent with the hypothesis of SLM [27] and SLM-r [28] that L2 experience affects how accurately learners perceive and produce L2 phonetic categories and their L2 perception and production remain malleable across the life-span as long as they are exposed to phonetic input differing from what they were exposed to previously in life. Mandarin learners of English were found to have difficulty producing native-like focus in English [29, 36]. In the present study, it is found that they also have difficulty perceiving English focus. This is consistent with the hypothesis of SLM-r that L2 speech perception and production coevolve and are closely associated with each other [28].

5. Conclusions

The present study investigated Chinese EFL learners' perception of English single and dual focus across focus location, intonation type, sentence length, and learners' English proficiency. Learners' perceptual accuracy was significantly impacted by focus location, intonation type and learners' English proficiency but not sentence length for both single and dual focus in English. Focus in statements was perceived better than that in questions. The interaction between focus and intonation led to high perceptual accuracy of focus located on sentence-final words and low accuracy of that located on non-final words in questions. Increased English proficiency resulted in the increased ability to perceive English focus accurately. The finding in the current study that learners have difficulty in perceiving L2 focus accurately may predict the difficulty for them in producing native-like prosodic focus in L2.

6. Acknowledgement

7. References

[1] S. Dik, "On the typology of focus phenomena," in *Perspectives on Functional Grammar*, T. Hoekstra, H. van der Hulst, and M. Moortgat, Eds. Berlin: De Gruyter, 1980, pp. 41–74.

[2] M. A. K. Halliday, *An Introduction to Functional Grammar*, 3rd ed., revised by C. Matthiessen. London: Hodder Arnold, 2004.

[3] W. E. Cooper, S. J. Eady, and P. R. Mueller, "Acoustical aspects of contrastive stress in question-answer contexts," *The Journal of the Acoustical Society of America*, vol. 77, no. 6, pp. 2142–2156, 1985.

[4] J. Pierrehumbert, "Prosody, intonation, and speech technology," in *Challenges in Natural Language Processing*, M. Bates and R. M. Weischedel, Eds. Cambridge: Cambridge University Press, 1993, pp. 257–280.

[5] Y. Xu, "Effects of tone and focus on the formation and alignment of F0 contours," *Journal of Phonetics*, vol. 27, no. 1, pp. 55–105, 1999.

[6] Y. Xu and C. X. Xu, "Phonetic realization of focus in English declarative intonation," *Journal of Phonetics*, vol. 33, no. 2, pp. 159–197, 2005.

[7] 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.

[8] Y. Xu, C. X. Xu, and X. Sun, "On the temporal domain of focus," in *Proc. Speech Prosody 2004*, Nara, Japan, Mar. 2004, pp. 81–84.

[9] W. L. Wu and L. Chung, "Post-focus compression in English-Cantonese bilingual speakers," in *Proc. ICPhS 2011*, Hong Kong, China, Aug. 2011, pp. 148–151.

[10] M. Vainio and J. Järvikivi, "Tonal features, intensity, and word order in the perception of prominence," *Journal of Phonetics*, vol. 34, no. 3, pp. 319–342, 2006.

[11] Y. C. Lee and Y. Xu, "Phonetic realization of contrastive focus in Korean," in *Proc. Speech Prosody 2010*, Chicago, US, May 2010, pp. 1–4.

[12] H. H. Rump and R. Collier, "Focus conditions and the prominence of pitch-accented syllables," *Language and Speech*, vol. 39, no. 1, pp. 1–17, 1996.

[13] M. Taheri-Ardali and Y. Xu, "Phonetic realization of prosodic focus in Persian," in *Proc. Speech Prosody 2012*, Shanghai, China, May 2012, pp. 326–329.

[14] Y. Chen, J. Zhang, B. Ye, and C. Zhou, "Prosodic realization of focus in Changchun Mandarin and Nanjing Mandarin," in *Proc. IALP 2019*, Shanghai, China, Nov. 2019, pp. 492–497.

[15] W. L. Wu and Y. Xu, "Prosodic focus in Hong Kong Cantonese without post-focus compression," in *Proc. Speech Prosody 2010*, Chicago, US, May 2010, pp. 1–4.

[16] 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.

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

[18] F. Kügler, "Post-focal compression as a prosodic cue for focus perception in Hindi," *Journal of South Asian Linguistics*, vol. 10, no. 2, pp. 38–59, 2020.

[19] E. Destruel and C. Féry, "Prominence in French dual focus," *Language and Speech*, vol. 64, no. 2, pp. 319–345, 2021.

[20] B. Wang and C. Féry, "Prosody of dual-focus in German: Interaction between focus and phrasing," *Language and Speech*, vol. 61, no. 2, pp. 303–333, 2018.

[21] S. J. Eady, W. E. Cooper, G. V. Klouda, P. R. Mueller, and D. W. Lotts, "Acoustical characteristics of sentential focus: Narrow vs. broad and single vs. dual focus environments," *Language and Speech*, vol. 29, no. 3, pp. 233–251, 1986.

[22] F. Liu, "Single vs. double focus in English statements and yes/no questions," in *Proc. Speech Prosody 2010*, Chicago, US, May 2010, pp. 1–4.

[23] B. Wang and C. Féry, "Dual-focus intonation in Standard Chinese," in *Proc. O-COCOSDA/CASLRE 2015*, Shanghai, China, Oct. 2015, pp. 101–106.

[24] 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.

[25] S. J. Eady and W. E. Cooper, "Speech intonation and focus location in matched statements and questions," *The Journal of the Acoustical Society of America*, vol. 80, no. 2, pp. 402–415, 1986.

[26] B. Wang, C. Li, Q. Wu, X. Zhang, B. F. Wang, and Y. Xu, "Production and perception of focus in PFC and non-PFC languages: Comparing Beijing Mandarin and Hainan Tsat," in *Proc. INTERSPEECH 2012*, Portland, US, Sept. 2012, pp. 663–666.

[27] J. E. Flege, "Second language speech learning: Theory, findings, and problems," in *Speech Perception and Linguistic Experience: Issues in Cross-language Research*, W. Strange, Ed. Timonium, MD: York Press, 1995, pp. 233–277.

[28] J. E. Flege and O-S Bohn, "The revised speech learning model (SLM-r)," in *Second Language Speech Learning: Theoretical and Empirical Progress*, R. Wayland, Ed. Cambridge: Cambridge University Press, 2021, pp 3–83.

[29] Y. Chen, "Post-focus compression in English by Mandarin learners," in *Proc. ICPhS 2015*, Glasgow, UK, Aug. 2015, pp. 1–5.

[30] Y. Chen, S. Guion-Anderson, and Y. Xu, "Post-focus compression in second language Mandarin," in *Proc. Speech Prosody 2012*, Shanghai, China, May 2012, pp. 410–413.

[31] T. Piske, I. R. A. MacKay, and J. E. Flege, "Factors affecting degree of foreign accent in an L2: A review," *Journal of Phonetics*, vol. 29, no. 2, pp. 191–215, 2001.

[32] T. Piske, "Implications of James E. Flege's research for the foreign language classroom," in *Language Experience in Second Language Speech Learning: In Honor of James Emil Flege*, O-S. Bohn and M. J. Munro, Eds. Amsterdam: John Benjamins Publishing, 2007, pp. 301–314.

[33] S. G. Guion, J. E. Flege, S. H. Liu, and G. H. Yeni-Komshian, "Age of learning effects on the duration of sentences produced in a second language," *Applied Psycholinguistics*, vol. 21, no. 2, pp. 205–228, 2000.

[34] J. E. Flege, D. Birdsong, E. Bialystok, M. Mack, H. Sung, and K. Tsukada, "Degree of foreign accent in English sentences produced by Korean children and adults," *Journal of Phonetics*, vol. 34, no. 2, pp. 153–175, 2006.

[35] G. Jia, W. Strange, Y. Wu, J. Collado, and Q. Guan, "Perception and production of English vowels by Mandarin speakers: Age-related differences vary with amount of L2 exposure," *The Journal of the Acoustical Society of America*, vol. 119, no. 2, pp. 1118–1130, 2006.

[36] J. Zhang, Y. Chen, and J. Cui, "Prosodic realization of focus in English by bidialectal Mandarin speakers," in *Proc. APSIPA ASC 2019*, Lanzhou, China, Nov. 2019, pp. 2042–2047.

[37] F. Liu, "Intonation systems of Mandarin and English: A functional approach," Ph.D. dissertation, University of Chicago, 2009.

[38] H. Wickham, *ggplot2: Elegant Graphics for Data Analysis*, 2nd ed., Switzerland: Springer International Publishing, 2016.

[39] Y. Xu, "The PENTA model of speech melody: Transmitting multiple communicative functions in parallel," in *Proc. From Sound to Sense*, Massachusetts, US, June 2004, pp. 91–96.

[40] Y. Xu, "Speech melody as articulatorily implemented communicative functions," *Speech Communication*, vol. 46, no. 3–4, pp. 220–251, 2005.

[41] Y. Xu, A. Lee, S. Prom-on, and F. Liu, "Explaining the PENTA model: A reply to Arvaniti and Ladd," *Phonology*, vol. 32, no. 3, pp. 505–535, 2015.

[42] Y. Xu, S. Prom-on, and F. Liu, "The PENTA model: Concepts, use and implications," in *Prosodic Theory and Practice*, J. Barnes and S. Shattuck-Hufnagel, Eds. Cambridge: The MIT Press, 2022, pp. 377–407.