*BUCLD 40 Proceedings*
*To be published in 2016 by Cascadilla Press*
*Rights forms signed by all authors*

# Prosodic Focus Marking in Minority L1 Bai-Children Learning Mandarin Chinese as L2

Zenghui Liu[*], Aoju Chen, and Hans Van de Velde

## 1. Introduction

An important function of prosody is to highlight new information in a sentence (i.e. focus). In languages that use prosody for marking focus, the prosodic cues involved differ from one language to another. For instance, some languages, such as English (Cooper et al. 1985), Dutch (Rump and Collier 1996; Chen 2009) and Mandarin (Xu 1999; Yang and Chen 2014a), not only expand the pitch range and lengthen the duration of the focal constituent, but also compress the pitch range and intensity of the post-focal constituent. However, other languages only exploit duration for the same purpose, such as Hong Kong Cantonese (Wu and Xu 2010), Tsat (Wang et al. 2012) and Bai (Liu et al. 2014).

Due to these observed cross-linguistic differences in the manner of marking focus prosodically, the acquisition of prosodic focus marking in L2 learners who acquire another language after early childhood (i.e. late bilinguals) and bilinguals who acquire two languages in early childhood (i.e. simultaneous bilinguals and sequential bilinguals) has recently received considerable attention (Chen and Guion-Anderson 2012; He et al. 2011; Nava and Zubizarreta 2008; Chen et al. 2014). It has been found that similarity between the learners' L1 and L2 did not necessarily lead to the attainment of native-like prosodic focus marking in late bilinguals' L2. For instance, He et al. (2011) found that Mandarin learners of Dutch did not acquire phonetic implementation of focus accentuation in Dutch, although Mandarin and Dutch mark focus in a similar manner: expanding the pitch range and the duration of the focal constituent and compressing the pitch range and duration of the post-focal constituent. Similarly, Chen and Guion-Anderson (2012) found that native English speakers

[*] Zenghui Liu, Utrecht University, l.z.h.liu@uu.nl; Aoju Chen, Utrecht University, aoju.chen@uu.nl; Hans Van de Velde, Fryske Akademy, HVandeVelde@fryske-akademy.nl

[**] We are grateful to all the bilingual participants from Jinhe Primary School and to the children's parents for their cooperation. We also thank Shanpeng Duan and teachers of Jinhe Primary School for their enormous support for the field help, Anqi Yang and Anna Sara Romøren for their feedback. This study is supported by a scholarship from the Chinese Scholarship Council to the first author and a VIDI grant (276-89-001) from the Netherlands Organisation for Scientific Research to the second author.

learning Mandarin as L2 did not show native-like patterns of in-focus changes in intensity and pitch, across all lexical tones.

Prosodic focus marking not only seems difficult to acquire for late bilinguals, but also for sequential bilinguals. For example, Wang et al. (2012) examined the Mandarin production of native speakers of Tsat, a tone language of the Austronesian language family spoken in Sanya, Hainan, China by the Utsuls. They found that Tsat-Mandarin sequential bilinguals did not mark focus prosodically in the same way as Mandarin monolinguals did. Specifically, Tsat-Mandarin sequential bilinguals did not expand the pitch range of the focal constituent in either their native language, Tsat, or in their second native language, Mandarin. However, they consistently lengthened the duration of the focal constituent in both languages. Chen et al. (2014) examined the prosodic focus marking produced by Min-Mandarin sequential bilinguals from different age groups in both languages. These speakers spoke Quanzhou Southern Min as their L1 and acquired Mandarin as their L2 in China. It was found that younger bilinguals could produce Beijing-Mandarin-like prosodic focus marking, but older bilinguals could not. Chen et al. (2014) suggested that more Beijing-Mandarin-like input and intensive training of Putonghua (Standard Mandarin) could explain the monolingual-like production by younger bilinguals. These previous studies of the acquisition of prosodic focus marking in bilinguals examined adult late bilinguals or adult sequential bilinguals. Therefore, their discussions mainly focused on the ultimate attainment of prosodic focus marking in late bilinguals or sequential bilinguals' L2.

Gut (2001) and Queen (1996; 2001; 2006; 2012) explored prosodic development in bilingual children by examining several aspects of prosody. Gut (2001) examined whether "language mixing" or cross-linguistic productions were in evidence in the developing intonational phonology of a German-English bilingual child. Her results from the examination of question intonation did not show the evidence of cross-linguistic production. Queen (1996; 2001; 2006; 2012) focused on bilingual acquisition in the context of a bilingual community at the group level. She was interested in two distinctive rises produced by Turkish-German bilinguals. The author concluded that these two terminal rises were distinctive in their acoustic, distributional and pragmatic characteristics, and that these rises were present in bilingual's intonational grammar and conventionalized (Queen 2012).

However, with regard to prosodic focus marking, there is little known about sequential bilingual children's language development path, although it is known that monolingual children learn to use prosody to mark focus in their L1 in stages (see Chen 2011; Romøren and Chen 2014; Yang and Chen 2014a; Yang and Chen 2014b). For example, Yang and Chen (2014b) reported that Mandarin-speaking children started to use duration to differentiate focus from non-focus at age four.

The question that arises is thus how sequential-bilingual children acquire prosodic focus-marking. We address this question by examining the production of 6- to 7-year-olds, 9- to 10-year olds and 11- to 12-year-olds with Bai as L1

and Mandarin as L2. Bai, a Tibetan-Burman tone language spoken by the Bai minority group in China, only exploits duration to mark focus prosodically (Liu et al. 2014). In contrast to Bai, Mandarin uses both pitch range and duration for this purpose (see Xu 1999; Yang and Chen 2014a).

More specifically, we address (1) whether child L2 speakers of Mandarin mark focus prosodically, i.e. the effect of focus; (2) whether child L2 speakers of Mandarin use prosodic cues to distinguish focus types that differ in the size of the focal constituent, i.e. the effect of size; (3) whether they distinguish contrastive focus from non-contrastive focus using prosodic cues, i.e. the effect of contrastivity; and (4) how their development progresses.

## 2. Methodology

### 2.1 Participants

Twenty-five child L2 speakers of Mandarin participated in our experiments, including eight 6- to 7-year old (4 girls, average age = 6;11, SD = 0;4), eight 9- to 10-year old (4 girls, average age = 9;2, SD = 0;4) and nine 12- to 13-year old (4 girls, average age = 12;5, SD = 0;6). All participants were native speakers of Southern Bai and acquired Mandarin as their second language beginning at about age six. The 6- to 7-year olds had just been enrolled at a local primary school at the time of testing. The 9- to 10-year olds already had three years of formal Mandarin education. At the time of testing, the 12- to 13-year olds had just started the sixth year of their Mandarin education. All the participants came from local Bai communities, and studied in Jinhe Primary School, Xizhou County, Dali Bai Autonomous Prefecture, China.

### 2.2 Experimental materials

A semi-spontaneous production task elicited SVO sentences with different focus structures, adapted from Chen (2011). The production experiment was conducted by a female (27-year-old) experimenter using Mandarin Chinese. It aimed to elicit SVO sentences in five focus conditions: narrow-focus on the subject NP in sentence-initial position (NF-i); narrow-focus on the verb in sentence-medial position (NF-m); narrow-focus on the object NP in sentence-final position (NF-f); broad focus (BF) and contrastive-focus on the verb in sentence-medial position (CF-m). The focus condition was set up by a WH-question or a statement from the experimenter, as illustrated in examples (1) to (5), where focused constituents appear in square brackets.

(1) Experimenter: Look! The book. There is also a pair of scissors. It looks like someone cuts the book. Who cuts the book?
Participant: [THE BEAR] cuts the book. (NF-i)

(2) Experimenter: Look! The bear and the book. It looks like the bear does something with the book. What does the bear do with the book?
Participant: The bear [CUTS] the book. (NF-m)

(3) Experimenter: Look! The bear, it uses a pair of scissors. It looks like the bear cuts something. What does the bear cut?
Participant: The bear cuts [THE BOOK]. (NF-f)

(4) Experimenter: Look! This picture is very blurry. I can't see anything clearly. What does the picture show?
Participant: [THE BEAR CUTS THE BOOK]. (BF)

(5) Experimenter: Look! The bear and the book. It looks like the bear does something with the book. I guess the bear throws the book.
Participant: The bear [CUTS] the book. (CF-m)

Eighty sentences were elicited from each participant. Among these eighty sentences, each focus condition (n = 5) was realized in sixteen SVO sentences. High level tone (tone 1), rising tone (tone 2), dipping tone (tone 3) and falling tone (tone 4) were included in subject noun phrase, verb and object noun. The realizations of the sentence-medial verbs were analysed and compared over the five focus conditions.

### 2.3. Data elicitation

The experiment was designed as a picture-matching game. In this picture-matching game, three piles of pictures were used: the experimenter and the participant each held a pile of pictures ordered in a certain sequence; the third pile of pictures were scattered around on a table. In the experimenter's pictures (the first pile), there was always something missing, like a subject, an action (verb) or an object. The participant's pictures (the second pile) all contained a complete event. The participant's task was to help the experimenter with sorting out pictures from her own pile and the third pile that went together. Here is a detailed example of a trial eliciting a target sentence in the NF-i condition: First, the experimenter took a picture (e.g. a book and a pair of scissors) from her own pile, drew the participant's attention to the picture and established what the picture was by saying, e.g. "*Look! The book. There is also a pair of scissors. It looks like someone cuts the book*." This was done to make sure that the entity in the picture was referentially given to the participant before the utterance of the question. Second, the experimenter asked a question about the picture (e.g. "*Who cuts the book*?"). Third, the participant took a complete picture from his or her pile and looked at it. The experimenter then repeated the question, followed by an answer from the participant (e.g. *"[THE BEAR] cuts the book*."). Fourth,

the experimenter found the picture containing the missing information in the third pile and paired it up with her own picture. The participants were explicitly instructed (1) to respond in full sentences and (2) not to show their own pictures to the experimenter. Prior to the picture-matching game proper, the experimenter conducted five practice trials with the participant to familiarize him or her with the game.

### 2.4. Recording procedure and equipment

All the experiments were conducted in a quiet room in Jinhe Primary School. Each experiment took 25-30 minutes per participant. Due to the limited concentration capacity of children, the experiment was divided into two sessions separated by a break, each consisting of 40 trials. The experiments were recorded using a portable ZOOM H1 digital recorder at a 44.1 KHz sampling rate and 16 bit accuracy. Each session was also video-taped.

### 2.5. Data selection and acoustic analysis

The target SVO sentences were selected and evaluated using Praat (Boersma 2001). Sentences that met the following criteria were selected for analysis: the sentence (1) was a response to the target question; (2) contained no self-corrections or hesitations; and (3) did not deviate from the target sentence in lexical choice and word order. In total, 1433 usable SVO target sentences were obtained from twenty-five speakers, from 2000 trials (72%).

## 3. Analysis and results

In order to investigate the effect of focus and focus types, we compared the word duration and pitch range of the target verb in various focus conditions. Specifically, to investigate the effect of focus, we compared phonetic measurements of the verbs between unfocused conditions and the focused condition, i.e. NF-m (verb is focused) vs. NF-i (verb is post-focal); and NF-m (verb is focused) vs. NF-f (verb is pre-focal). To investigate the effect of focus type that differs in the size of the focal constituent, we compared narrow focus on the verb (NF-m) with the broad focus condition (BF). To investigate the effect of contrastivity, we compared contrastive focus on the verb (CF-m) with non-contrastive focus on the verb (NF-m).

Statistical analyses were conducted using mixed effect modelling in R (R Core Team 2014). More specifically, package "lme4" (Bates et al. 2014) and "lmerTest" (Kuznetsova et al. 2014) were used. In all models, *Focus Condition*, *Verb Tone* and *Age* were included as fixed factors, while *Speaker* and *Sentence* were included as random factors. In each of the four pair-wise comparisons listed above, *Focus Condition* had two levels, the two focus conditions of

interest; *Verb Tone* had four levels which referred to the four lexical tones of the target verbs; and *Age* had three levels which referred to the three age groups of the participants. The experimental design contained one representation for each lexical tone, i.e. only one word containing each specific lexical tone was included. Dependent variables were the duration and the pitch range of the target verbs.

Models were constructed and evaluated in a stepwise fashion, starting from the model with only random factors. When building the models, only factors and interactions that significantly improved the previous model were retained in subsequent models. The improvement of the model fit at each step was assessed by the difference in -2LL (log likelihood), i.e. a statistically significant difference between two models was an indication of a significant effect of the added fixed factor. We excluded the candidate models that did not lead to significant improvements over the previous model to arrive at the best fitting model. Using this procedure we could assess the effect of the factors listed, as well as their interactions.

### 3.1 Effect of focus: narrow focus vs. non-focus

#### 3.1.1. Duration

The duration data obtained from the verbs in the NF-m and NF-i conditions showed that the verbs were on average 18.5 ms longer when focused (NF-m) than when not focused and following a focused constituent (NF-i). Mixed-effect modeling was used to assess the effect of *Focus Condition* on the duration of the verbs, as described above. It revealed a main effect of *Focus Condition* ($\chi^2$ (1) = 6.2818, $p < .05$). The best fitting model was the one containing main effects of *Focus Condition*, *Verb Tone* and *Age*, in addition to an interaction between *Verb Tone* and *Age*. However, neither a two-way interaction between *Focus Condition* and *Verb Tone* ($p = .42$) nor a two-way interaction between *Focus Condition* and *Age* ($p = .16$) was found. This suggests that the all the age groups lengthened the duration of the focal constituent in all the lexical tones in comparison to the post-focal constituent, but that each verb was produced with different duration by different age groups.

The duration data obtained from the verbs in the NF-m and NF-f conditions showed that the verbs were on average 14.6 ms longer when focused (NF-m) than when not focused and preceding a focused constituent (NF-f). Mixed-effect modelling revealed a main effect of *Focus Condition* ($\chi^2$ (1) = 5.58, $p < .05$). The best fitting model was the one containing main effects of *Focus Condition*, *Verb Tone* and *Age*, in addition to an interaction between *Verb Tone* and *Age*. Similarly to the previous comparison, neither a two-way interaction between *Focus Condition* and *Verb Tone* ($p = .43$), nor a two-way interaction between *Focus Condition* and *Age* ($p = .41$) was found. This suggests that the all the age groups lengthened the focal constituent in all the lexical tones in comparison to the pre-focal constituent, although the duration of each verb was different in

different age groups. The use of duration for distinguishing the focal constituent from the post-focal and pre-focal constituent is shown in **Figure 1**.

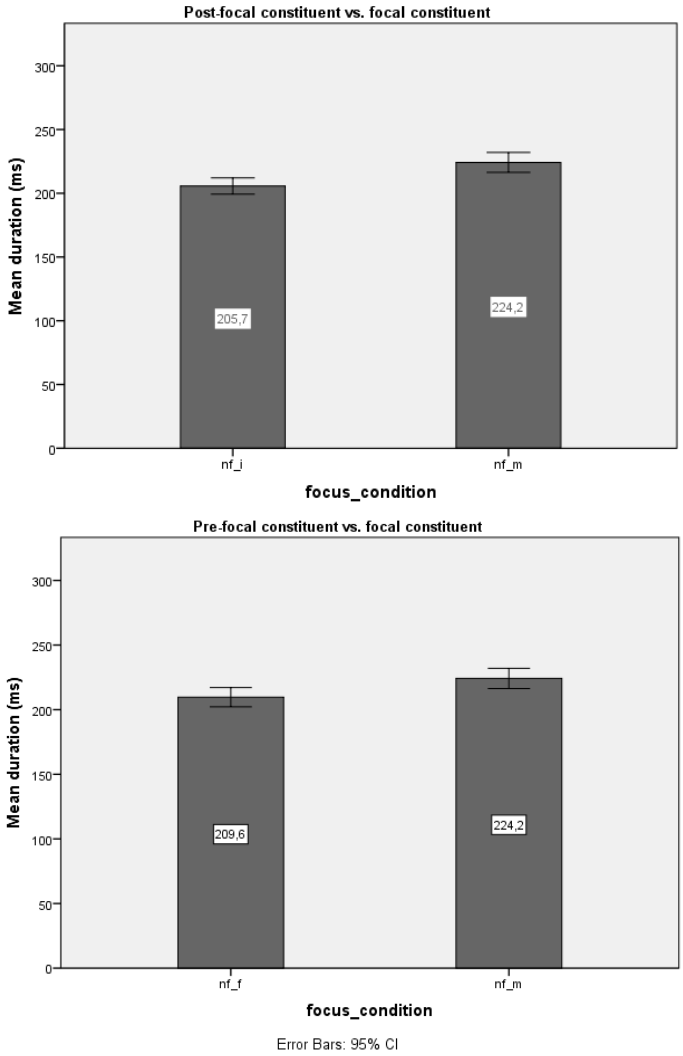



**Figure 1: Mean duration (in ms) of the post-focal constituent vs. focal constituent (upper panel) and pre-focal constituent vs. focal constituent (lower panel) across all four lexical tones**

### 3.1.2 Pitch range

With regard to the comparison between the focal constituent (NF-m condition) and the post-focal constituent (NF-i condition), the best fitting model was the model with a three-way interaction between *Focus Condition*, *Verb Tone* and *Age* and these three fixed effects ($\chi^2(6) = 13$, $p < .05$). This three-way interaction suggests that the use of pitch range as a prosodic cue for differentiating the focal constituent from the post-focal constituent was different in different lexical tones in different age groups. In order to scrutinize the difference between different lexical tones in different age groups, we reset the intercept of each condition when necessary and then re-ran the model. The comparisons we were interested in were the differences between focus conditions (NF-i vs. NF-m) per age group and per tone. For 6- to 7-year olds, the pitch range of the verbs was on average 15.9 Hz larger when focused (NF-m) than when not focused and following a focused constituent (NF-i) in tone 4 ($p < .05$), but the pitch range of the verbs was on average 16.2 Hz smaller when focused (NF-m) than when not focused and following a focused constituent (NF-i) in tone 3 ($p < .05$). For other age groups and other lexical tones, no statistically significant differences were found in the comparison between the focal constituents and the post-focal constituents. The use of pitch range for distinguishing the focal constituents from the post-focal constituents in four lexical tones by 6- to 7-year olds is shown in **Figure 2**.

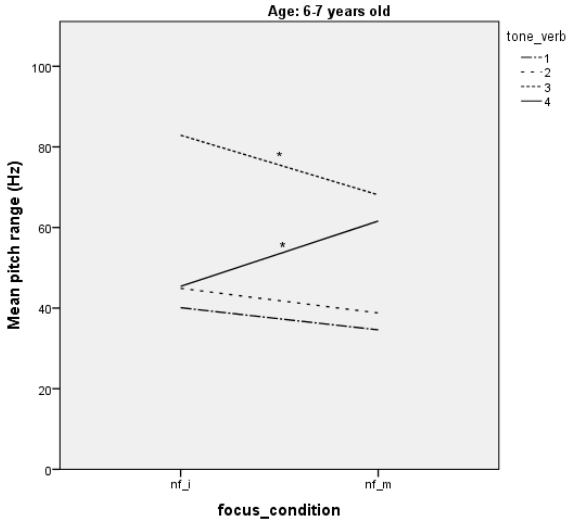



**Figure 2: Mean pitch range (in Hz) of post-focal constituent vs. focal constituent in four lexical tones in 6- to 7-year olds**

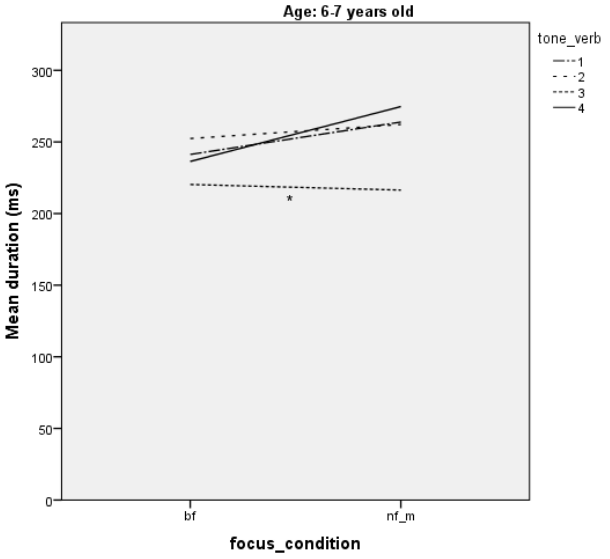
With regard to the comparison between the focal constituent (Nf-m) and the pre-focal constituent (NF-f), the best fitting model was the model with a two-way interaction between *Verb Tone* and *Age*. There was thus no evidence that

the sequential bilinguals varied the pitch range in their Mandarin to distinguish the focal constituent from the pre-focal constituent.

### 3.2. Effect of size: narrow focus vs. broad focus

#### 3.2.1. Duration

To test for an effect of focus type differing in size on the duration of target verb, mixed effect modelling as described above was conducted. The best fitting model we obtained was the model with a three-way interaction between *Focus Condition*, *Verb Tone* and *Age* and these three fixed effects ($\chi^2(6) = 13.12$, $p < .05$). This three-way interaction suggested that the use of duration as a prosodic cue for differentiating the focal constituent in broad focus condition from narrow focus condition was different in different lexical tones by different age groups. For 6- to 7-year olds, the duration of the verbs were on average 49.8 ms shorter when in the narrow focus condition (NF-m) than when in the broad focus condition (BF) in tone 3 ($p < .01$). In addition, 9- to 10-year olds produced the target verb on average 27 ms shorter in the narrow focus condition (NF-m) in comparison to the broad focus condition (BF) in tone 4 ($p < .05$). For other age groups and other lexical tones, no statistically significant differences were found in the comparison between narrow focus and broad focus. The use of duration to distinguish the focus types differing in size in four lexical tones by the 6- to 7-year olds and the 9- to 10-year olds is shown in **Figure 3.**




.

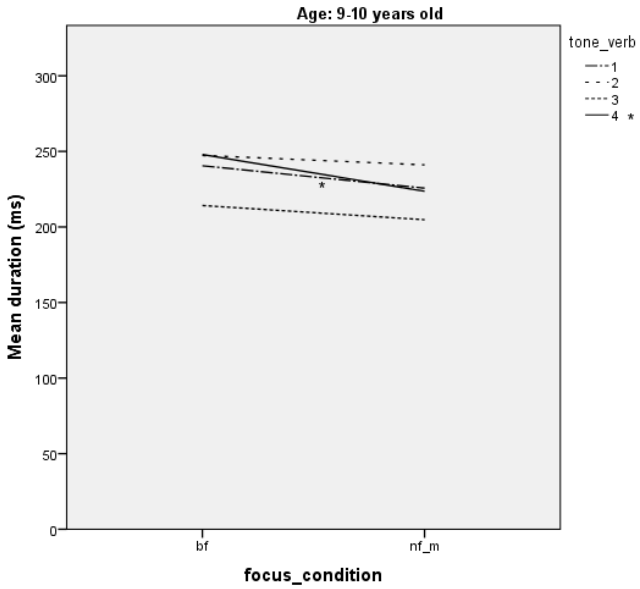



**Figure 3: Mean duration (in ms) of the focal constituent in broad focus condition vs. narrow focus condition in four lexical tone in 6- to 7-year olds (upper panel) and in 9- to 10-year olds (lower panel).**

### 3.2.2. Pitch range

With regard to the effect of size on the pitch range of the focal constituent in the broad focus condition and narrow focus condition, we compared the BF condition with NF-m condition. The best fitting model we obtained was the model with a two-way interaction between *Verb Tone* and *Age*. There was thus no evidence that our sequential bilinguals varied the pitch range in their Mandarin to distinguish the focal constituent in the broad focus condition from narrow focus condition.

## 3.3. Effect of contrastivity

### 3.3.1. Duration

With regard to the effect of focus types differing in contrastivity on the duration of target verb, mixed effect modelling as described above was conducted. The best fitting model we obtained was the model with a three-way interaction between *Focus Condition*, *Verb Tone* and *Age*, and these three fixed effects ($\chi^2(6) = 21.14$, $p < .01$). This three-way interaction suggests that the use of duration as a prosodic cue for differentiating the focal constituent in the

contrastive focus condition from the non-contrastive focus condition was different in different lexical tones in different age groups. For 6- to 7-year olds, the duration of the verbs was on average 37.6 ms longer when in the contrastive focus condition (CF-m) than when in the non-contrastive focus condition (NF-m) in tone 3 ($p < .01$). For other age groups and other lexical tones, no statistically significant differences were found in the comparison between contrastive focus and non-contrastive focus. The use of duration for distinguishing the focus types differing in contrastivity in four lexical tones by 6- to 7- year olds is shown in **Figure 4**.

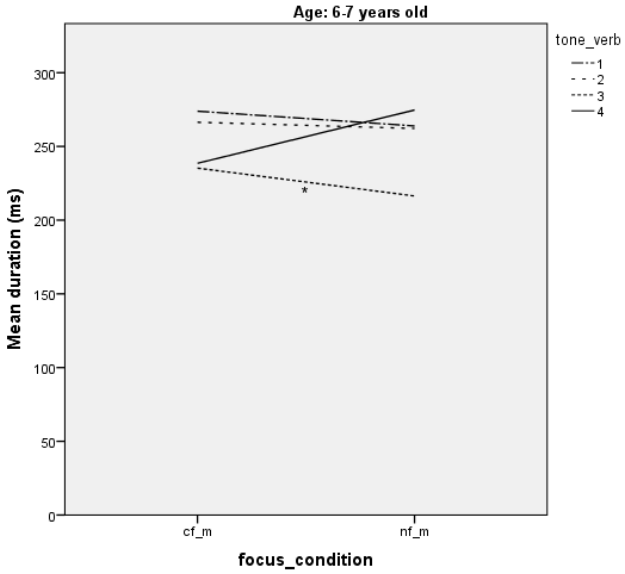



**Figure 4: Mean duration (in ms) of contrastive focal constituent vs. non-contrastive focal constituent in four lexical tones in 6- to 7-year olds**

### 3.3.2. Pitch range

When examining the pitch range of the focal constituent in the contrastive focus condition (CF-m) and non-contrastive focus condition (NF-m), the best fitting model was the model with a two-way interaction between *Verb Tone* and *Age*. There was thus no evidence that the sequential bilinguals varied the pitch range in their Mandarin to distinguish contrastive focal constituents from non-contrastive focal constituents.

## 4. Conclusion and discussion

The present study examines prosodic focus marking produced by sequential bilingual children whose L1 is Bai and who started their formal Mandarin education at about the age of six. Our results show that this population of Bai-Mandarin sequential bilinguals mark focus prosodically in their Mandarin by using duration, but their production is not completely Mandarin monolingual-like.

With regard to our first research question, whether child L2 speakers of Mandarin mark focus prosodically, our results show that Bai-Mandarin sequential bilingual children of all the age groups tested consistently use duration to mark focus prosodically. More specifically, Bai-Mandarin sequential bilinguals increase the duration of the focal constituent in comparison to pre-focal and post-focal constituents in their Mandarin production. In terms of using pitch range as a prosodic cue for marking focus, different age groups behave differently. The use of pitch range to mark focus prosodically interacts with the lexical tone of the focal constituent for 6- to 7-year olds. The 6- to 7- year olds seem to differentiate the focal constituent from the post-focal constituent only in tone 3 and tone 4, but in different directions. The 6- to 7-year olds compress the pitch range of the focal constituent in tone 3 rather than expand the pitch range as a monolingual Mandarin speaker would do (Xu 1999; Yang and Chen 2014a). However there is no evidence that 6- to 7-year olds make such a distinction in other lexical tones, or that other age groups make this distinction at all.

With regard to our second research question, whether child L2 speakers of Mandarin use prosodic cues to distinguish focus types that differ in the size of the focal constituent, our results show an interaction between the use of duration as a prosodic cue for distinguishing focus types that differ in size and the lexical tone of the focal constituent in different age groups. We find that only 6- to 7-year olds make a duration distinction between different focus types differing in size in tone 3, and 9- to 10-year olds make the same distinction in tone 4. However, the distinctions that they make are opposite to monolingual Mandarin production (Xu 1999; Yang and Chen 2014a). As for the use of pitch range to mark focus types that differ in size prosodically, none of the age groups showed that capability in any lexical tones. Therefore, the use of duration and pitch range as prosodic cues for distinguishing focus types that differ in the size of the focal constituent (i.e. the effect of size) is not Mandarin-monolingual like.

With regard to our third research question, whether child L2 speakers of Mandarin distinguish contrastive focus from non-contrastive focus using prosodic cues, our results show that pitch is used as a prosodic cue to make a distinction between contrastive focus and non-contrastive focus by sequential bilinguals in their Mandarin. However, the 6- to 7-year olds in the present study expand the pitch range of the contrastive-focal constituent only in tone 3 in comparison to non-contrastive focal constituent. As for other lexical tones and other age groups, there is no evidence of Bai-Mandarin sequential bilingual children's ability to mark contrastive focus prosodically.

To our knowledge, the present study is the first attempt to reveal the prosodic developmental path of sequential bilinguals in terms of prosodic focus

marking. Our results show that the sequential bilinguals' acquisition of prosodic focus marking proceeds by stages. Bai-Mandarin sequential bilingual children seem to be quite good at using duration as a prosodic cue for distinguishing the focal constituent from the non-focal constituent. However, the use of duration for differentiating focus types differing in size and contrastivity is not consistently found in all the age groups or consistently across lexical tones. Our results suggest two accounts for the early mastery of the use of duration in L2 for marking focus. One is that duration as a prosodic cue may be easier for learners. The other is that the sequential bilingual children benefitted from positive L1 transfer, as their native language, Bai, also exploits duration as a prosodic cue for the same purpose. It is worth mentioning that the 6- to 7-year olds in our study are sequential bilinguals who had first enrolled at the local primary school one week prior to testing. It is therefore unlikely that they had started to pick up the use of duration for marking focus from their new L2, Mandarin. We therefore suggest that the early mastery of the use of duration for marking focus by Bai-Mandarin sequential bilinguals arises from positive L1 transfer.

In addition, the use of pitch range for marking focus also differs in different age groups in different lexical tones. Interestingly, the interactions we found between focus and lexical tones are present in the 6- to 7-year olds and 9- to 10-year olds, but absent in 11- to 12-year olds. The complex interactions provide evidence that sequential bilinguals struggle to use the same prosodic parameters, i.e. pitch range and duration, for the dual purposes of making lexical distinctions (i.e. lexical tones) and encoding information structure (i.e. focus marking) at the same time.

## References


Bates, Douglas, Martin Maechler, Ben Bolker and Steven Walker (2014). "lme4: Linear mixed-effects models using Eigen and S4." ArXiv e-print; submitted to *Journal of Statistical Software*, <URL: http://arxiv.org/abs/1406.5823>.

Boersma, Paul (2001). "Praat, A System for Doing Phonetics by Computer." *Glot International* 5:9/10, 341-345.

Chen, Aoju. 2011. "The Developmental Path to Phonological Focus-Marking in Dutch." *In Prosodic Categories: Production, Perception and Comprehension*, 93-109: Springer.

Chen, Aoju. 2009. "The Phonetics of Sentence-Initial Topic and Focus in Adult and Child Dutch." *Phonetics and Phonology: Interactions and Interrelations*: 91-106.

Chen, Ying, Yi Xu, and Susan Guion-Anderson. 2014. "Prosodic Realization of Focus in Bilingual Production of Southern Min and Mandarin." *Phonetica* 71 (4): 249-270.

Chen, Ying and Susan Guion-Anderson. 2012. "Prosodic Realization of Focus in Mandarin by Advanced American Learners of Chinese." *The Journal of the Acoustical Society of America* 131 (4): 3234-3234.

Cooper, William E., Stephen J. Eady, and Pamela R. Mueller. 1985. "Acoustical Aspects of Contrastive Stress in Question–answer Contexts." *The Journal of the Acoustical Society of America* 77 (6): 2142-2156.

Gut, Ulrike. 2001. "Cross-Linguistic Structures in the Acquisition of Intonational Phonology by German-English Bilingual Children." *Crosslinguistic Structures in Simultaneous Bilingualism*: 201-226.

He, Xuliang, Judith Hanssen, Vincent J. van Heuven, and Carlos Gussenhoven. 2011. "Phonetic Implementation Must Be Learnt: Native Versus Chinese Realization of Focus Accent in Dutch." *In Proceedings of the XVIIth International Congress of Phonetic Sciences* (pp. 843-846)

Kuznetsova, Alexandra, Per Bruun Brockhoff, and Rune Haubo Bojesen Christensen. "lmerTest: Tests for Random and Fixed Effects for Linear Mixed Effect Models (lmer objects of lme4 package)." *R package version 2*, no. 6 (2013).

Liu, Zenghui, Aoju Chen, and Hans Van de Velde. 2014. "Prosodic Focus Marking in Bai." *In Proceedings of Speech Prosody 2014*.

Nava, Emily and Maria Luisa Zubizarreta. 2008. "Prosodic Transfer in L2 Speech: Evidence from Phrasal Prominence and Rhythm." *In Proceedings of Speech Prosody 2008*: 335-338.

Queen, Robin Michelle. 2001. "Bilingual Intonation Patterns: Evidence of Language Change from Turkish-German Bilingual Children." *Language* in Society 30 (01): 55-80.

Queen, Robin Michelle. 2006. "Phrase-Final Intonation in Narratives Told by Turkish-German Bilinguals." *International Journal of Bilingualism* 10 (2): 153-178.

Queen, Robin Michelle. 1996. "Intonation in Contact: A Study of Turkish-German Bilingual Intonation Patterns". Unpublished Disseration. University of Texas, Austin. University of Texas at Austin.

Queen, Robin. 2012. "Turkish-German Bilinguals and their Intonation: Triangulating Evidence about Contact-Induced Language Change." *Language* 88 (4): 791-816.

R Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Romøren, Anna Sara H. and Aoju Chen. 2014. "Accentuation, Pitch and Duration as Cues to Focus in Dutch 4-to 5-Year-Olds." *In Proceedings of BUCLD 2013*.

Rump, Hans H. and Rene Collier. 1996. "Focus Conditions and the Prominence of Pitch-Accented Syllables." *Language and Speech* 39 (1): 1-17.

Wang, Bei, Chenxia Li, Qian Wu, Xiaxia Zhang, Baofeng Wang, and Yi Xu. 2012. "Production and Perception of Focus in PFC and Non-PFC Languages: Comparing Beijing Mandarin and Hainan Tsat." *In Proceedings of Interspeech 2012*

Wu, Wing Li and Yi Xu. 2010. "Prosodic Focus in Hong Kong Cantonese without Post-Focus Compression." *In Proceedings of Speech Prosody 2010*.

Xu, Yi. 1999. "Effects of Tone and Focus on the Formation and Alignment of F 0 Contours." *Journal of Phonetics* 27 (1): 55-105.

Yang, Anqi and Aoju Chen. 2014a. "Prosodic Focus Marking in Child and Adult Mandarin Chinese." *In the 4th International Symposium on Tonal Aspects of Language* (pp. 54-58).

Yang, Anqi and Aoju Chen. 2014b. "Prosodic Focus-Marking in Chinese Four-and Eight-Year-Olds." *In Proceedings of Speech Prosody. 2014*