Prosodic Realization of Focus in Bilingual Production of Southern Min and Mandarin

Ying Chen[a] Yi Xu[b] Susan Guion-Anderson[a]

[a] Department of Linguistics, University of Oregon, Eugene, OR 97403, USA

[b] Department of Speech, Hearing and Phonetic Sciences, Division of Psychology and Language Sciences, University College London, London, UK

Short title: Prosodic Focus in Southern Min and Mandarin

Key words: Prosodic focus, post-focus compression, bilingual, Southern Min, Mandarin

Ying Chen

Department of Linguistics

University of Oregon

Eugene, OR 97403

USA

Tel: 1- 541-346-3906

Fax: 1-541-346-5961

E-mail: ychen12@uoregon.edu

Abstract

Previously post-focus compression (PFC) — the lowering of F0 and intensity of post-focal words to below those of the same words in identical sentences with neutral-focus — was found in Beijing Mandarin but not in Taiwan Southern Min and Taiwan Mandarin. This study investigated whether the presence of PFC would vary with age and language use of societal bilinguals of Southern Min and Mandarin. Three groups of bilingual speakers of Quanzhou Southern Min and Mandarin, age around 20, 40 and 60, were examined for their prosodic realization of focus. All the speakers acquired Southern Min first, followed by Mandarin in childhood, but the younger speakers used more Mandarin than the older speakers. Comparisons of duration, intensity and F0 in focused, pre-focus and post-focus words indicated that all groups produced Taiwan-like focus, i.e., without PFC, in Southern Min, but the youngest group produced Beijing-like PFC in Mandarin. These findings reveal that increased language experience, such as greater amount of L2 use, correlates with increased ability to produce native-like PFC in L2, suggesting that PFC can be used as an indicator in assessing L2 speech acquisition.

1. Introduction

Prosodic focus is a phonetic means of highlighting part of an utterance against the rest of the constituents [Bolinger, 1972; Gussenhoven, 1983; Lambrecht, 1994; Selkirk, 1995, 2006][1]. Early studies have found that the focused constituent is acoustically marked by variations in pitch, intensity and duration in English [Pierrehumbert, 1980, 1993; Cooper, Eady and Mueller, 1985; Eady and Cooper, 1986]. More recent studies have shown that in many languages, not only does a focused constituent itself go through changes in duration, F_0 and intensity under focus but also there is post-focus compression (PFC) of F_0 and intensity, relative to a neutral focus sentence [English: Cooper, Eady and Mueller, 1985; Eady and Cooper, 1986; Xu and Xu, 2005. Finnish: Vainio and Järvikivi, 2007. Dutch: Hanssen, Peters and Gussenhoven, 2008. Hindi: Patil et al., 2008. Japanese: Kubozono, 2009. Korean: Lee and Xu, 2010]. It has also been found that PFC is not unique to non-tone languages, as it also occurs in tone languages like Mandarin [Jin, 1996; Xu, 1999] and the Nanchang dialect of Chinese [Wang, Wang and Kadir, 2011]. More interestingly, there is also emerging evidence that PFC is absent in many other language, some tonal and some non-tonal [Xu, 2011; Xu, Chen and Wang, 2012], indicating that it is not universal. Further, tone and focus have been found to be largely independent of each other. That is, in some languages, such as Mandarin [Xu, 1999] and Nanchang [Wang, Wang and Kadir, 2011], PFC occurs regardless of the tones, while in other languages, such as Cantonese [Wu and Xu, 1999] and Yi [Wang, Wang and Kadir,

[1] Féry [2013] pointed out that not all languages were observed to code focus by using prosody.

2011], it is simply absent, also regardless of the tones.

The finding of differential cross-linguistic distribution of PFC has opened up a new avenue for the investigation of second language acquisition and bilingualism. Xu, Chen and Wang [2012] examined focus realization in four groups of speakers in Taiwan and Beijing: monolinguals of Taiwan Southern Min (Taiwanese), monolinguals of Taiwan Mandarin, bilinguals of Taiwan Southern Min and Taiwan Mandarin, and monolinguals of Beijing Mandarin. Although the changes in F_0, intensity and duration were similar in focused words in the production of both languages in the four groups, only monolingual Beijing Mandarin speakers produced PFC of F_0 and intensity. PFC was absent in Taiwan Southern Min produced by both Taiwan Southern Min monolinguals and Taiwan Southern Min-Mandarin bilinguals, and even in Taiwan Mandarin produced by Taiwan Mandarin monolinguals and Taiwan Southern Min-Mandarin bilinguals. That study has therefore provided the first explicit evidence of two phenomena: (a) the presence of PFC can differ across languages/dialects that are closely related, since Beijing Mandarin, Taiwan Mandarin and Taiwanese all belong to the Chinese language family but PFC is present only in Beijing Mandarin; (b) the presence of PFC can be altered by language contact, since Taiwan Mandarin seems to have lost it due to contact with Taiwanese, mainly through bilingualism.

These findings suggest that the use of PFC may correlate with second language proficiency, especially in the case of speakers of a non-PFC language learning a PFC language. The current study explores this possibility by examining the case of societal bilingualism in Quanzhou, China.

Since the National Language Movement in the mid 1940s, Mandarin has been promoted and spread as a national language in Taiwan [Li and Lee, 2006]. Taiwan has since become a large speech community of Taiwan Southern Min-Mandarin bilinguals. Due to the political separation of Taiwan and mainland China and contact with Taiwan Southern Min since 1949, the standard Mandarin in Taiwan has deviated from Beijing Mandarin, and has become diversified also due to different degrees of contact with Taiwan Southern Min [Liao, 2008]. Phonological interference or transfer as a result of language contact and bilingualism has been observed for many contact situations [Thomason and Kaufman, 1988; Sankoff, 2001]. According to Van Coetsem's [1988] "phonological imposition" theory of language contact, Mandarin, as a recipient language in Taiwan, has acquired some phonological/phonetic features from the source language—Taiwan Southern Min. From the viewpoint of second language acquisition, Taiwan Southern Min-Mandarin bilinguals typically acquire Southern Min earlier than Mandarin. They learn and use Southern Min as L1 at home and Mandarin as L2 at school [Huang and Fon, 2007]. Therefore, their Mandarin phonological system is influenced by their Southern Min phonological system [Kubler, 1985]. The findings of Xu et al. [2012] indicate that such influence can be manifested in not only segmental but also prosodic aspects of speech.

Diglossia in Quanzhou, a city where Southern Min has been spoken for over 1,500 years in mainland China and one of the cities from which most of the population of Taiwan is derived, is similar to diglossia in Taiwan. The residents there speak not only Quanzhou Southern Min, but also Beijing Mandarin (Putonghua) as required by the government policy of National Popularization of Putonghua. Since the implementation of the policy in the 1950s, local residents in Quanzhou have been educated and immersed in Beijing Mandarin and become Southern Min-Mandarin bilinguals. People around 60 years of age are considered the first generation of Southern Min-Mandarin bilinguals in Quanzhou. However, their daily use of Mandarin is quite limited. After several decades of popularization of Putonghua, and with continuous increase in language contact, the younger generation uses more Mandarin than the older generations. Additionally, most of the younger speakers, unlike the older ones, receive preschool education in Mandarin; so their age of learning (AOL) L2 Mandarin tends to be slightly earlier than the older speakers. Language experience, whose factors includes AOL, the amount of L1/L2 use, length of residence in the L2 speaking environment, the quality of L2 input, etc. [Piske et al., 2001; Piske, 2007], therefore varies by age in Quanzhou Southern Min-Mandarin bilinguals. This provides an opportunity to determine whether the lack of PFC in a language, as in the case of Mandarin spoken by Southern Min-Mandarin speakers [Xu et al., 2012], is related to language experience, and especially to language use. The current study makes use of this opportunity.

Previous studies of the effects of language experience on bilingual speech production have mostly focused on immigrant bilinguals, namely those who start learning L2 after they have immigrated to the L2 speaking environment or the second generation of immigrants who speak the local language and also their parents' L1. However, the factors found to be relevant to immigrant bilinguals have also been confirmed in studies of societal bilinguals, namely those who speak two languages that are both required and used in the community or society [Peng, 1993; Guion, Flege and Loftin, 2000; Guion, 2003]. AOL has been found to be one of the most important factors to affect L2 speech production [Piske, MacKay and Flege, 2001]. Early bilinguals tend to produce more native-like vowels [Munro, Flege and MacKay, 1996; Flege, MacKay and Meador, 1999] and consonants [Flege, Munro and MacKay, 1995] than later bilinguals. AOL in immigrant bilinguals also influences the interaction between L1 and L2 vowels [Baker and Trofimovich, 2005] and consonants [Kang and Guion, 2006]. At the suprasegmental level, AOL affects stress, speech rate and intonation patterns in second language speech [Guion, Flege, Liu and Yeni-Komshian, 2000; Lee, Guion and Harada, 2006; Huang and Jun, 2011]. In addition to AOL, Flege, Frieda and Nozawa [1997] found that the amount of L1 use is also important. Italian immigrants in Canada with more L1 Italian use were rated as having a stronger foreign accent in L2 English than those with less Italian use. In regard to societal bilingualism, Guion, Flege and Loftin [2000] found that subjects with more L1 use of Quichua (a Quechuan language spoken in Otavalo, Ecuador) had stronger

foreign accent in L2 Spanish than those with less Quichua use; however, the degree of foreign accent in L1 Quichua was not significantly affected by the amount of Quichua use. These two studies therefore indicate that the amount of L1 use affects the degree of foreign accent in L2: the greater the L1 use, the greater the negative effect on L2 production.

There is also evidence of influences of L2 on L1 in social bilinguals: vowel production is linked to AOL [Guion, 2003] and consonant production is linked to L2 proficiency [Peng, 1993]. In Quichua-Spanish societal bilingualism, Guion [2003] found that simultaneous bilinguals were able to produce monolingual-like vowels in both languages, and that early and some mid bilinguals were able to produce native-like L2 Spanish vowels while late bilinguals were not. The Quichua vowels were produced higher by bilinguals who had acquired Spanish vowels than those who had not. In the societal bilingualism of Southern-Min and Mandarin, Peng [1993] found that the Mandarin production of /f/ and /x/ varied in the spectrum from native-like Mandarin /f/ and /x/ to Southern Min /h^w/ and / h/ with a positive relation to the degree of Mandarin proficiency. Subjects with the highest proficiency in Mandarin showed some interference from Mandarin [x] on their Southern Min production. These two studies demonstrate the mutual influence of L1 and L2 system in societal bilingualism and especially the effect of high proficient L2 on the L1 production.

More specifically related to the second language acquisition of prosodic focus, McGory [1997] found that nonnative English speakers did not vary F_0 of stressed syllables according to intonational context as did native speakers, and produced higher F_0 in stressed than in unstressed syllables in both focused and unfocused words. The degree of this tendency varied by subjects' L1 (Korean vs. Mandarin) and L2 English experience. Nava and Zubizarreta [2008] found that two out of ten Spanish learners of English acquired both what they referred to as the nuclear accent rule and the post-nuclear anaphoric de-accenting rule in L2 English; two acquired post-nuclear Anaphoric De-accenting but not the nuclear accent rule; and six did not acquire either rule. He et al. [2011] found that, compared to native Dutch speakers, Chinese learners of Dutch did not show a regular pattern of prosodic features in broad, narrow or corrective focus. Intriguingly, the Chinese learners' production of Dutch did not show differences between the groups with high and low proficiency in Dutch. That is, their uncertain production of prosodic focus in Dutch could not be attributed to their L2 proficiency. Wu and Chung [2011] found that eight out of ten English-Cantonese simultaneous bilingual subjects had PFC in their English but not in their Cantonese. They concluded that PFC is not easily transferred to a non-PFC language even through simultaneous bilingualism, the most intimate form of language contact.

In summary, previous research has shown that L2 speech production is influenced by language experience, and especially by age of L2 learning and amount of L1 or L2 use, in both segmental and prosody domains. Meanwhile, the recent finding of cross-linguistic

variation in the presence/absence of PFC suggests that PFC may also vary as a function of language experience in cases where L1 and L2 differ in prosodic focus realization. The current study is the first systematic effort to explore this possibility. The aim is to not only examine possible variation of PFC as a function of amount of L2 use, but also determine if there is a lack of influence of L2 on the production of prosodic focus in L1, as has been found in Xu et al. [2012] and Wu and Chung [2011]. The following specific research questions will be investigated: (1) Are there expansions of F_0, intensity and duration of focused words in both Southern Min and Mandarin by Quanzhou bilingual speakers? (2) Does PFC of F_0 and intensity occur in Quanzhou bilinguals' production of L1 Southern Min and L2 Mandarin? (3) Do different age groups produce different patterns of prosodic focus in both languages? Is there an intermediate pattern of prosodic focus in the bilingual production between the older and the younger generations? (4) Is there any reverse influence of L2 Mandarin on the prosodic patterns of focus in L1 Southern Min? Questions 2 and 4 are especially interesting because their outcomes could go either way based on previous findings. Two 2-way predictions are therefore conceivable:

For question 2: Prediction 1a — There is no PFC in L2 Mandarin regardless of language experience, just as found in Taiwan;

Prediction 1b — PFC occurs in L2 Mandarin when there is sufficient L2 experience.

For question 4: Prediction 2a — PFC occurs in L1 Southern Min, especially when L2 use becomes dominant, in line with previous findings of L2 to L1 transfer;

Prediction 2b — PFC does not occur in L1 Southern Min, regardless of L2 experience.

2. Methods

2.1 Participants

Three age groups of Quanzhou Southern Min-Mandarin bilinguals participated in this study: younger, mid-age and older. The younger speakers were between 18 and 21 years of age, the mid-age speakers between 35 and 43 years, and the older speakers between 55 and 64 years. There were four males and four females in each group. All participants were born and raised in Quanzhou City, Fujian Province, mainland China. Participants in the younger group had always lived in Quanzhou and were students at

Quanzhou Normal University at the time of testing. Participants in the older and mid-age groups had never lived longer than three months out of Quanzhou. All three groups had learned Southern Min first, followed by Mandarin in childhood. However, the amount of Mandarin use and age of learning Mandarin varied among the age groups. All the participants reported having normal hearing and speaking no Chinese languages other than Southern Min and Mandarin.

The language experience of L2 Mandarin was determined by participants' responses to a language background questionnaire (LBQ). The LBQ requested participants to report their chronological age, the age at which they learned Southern Min and Mandarin, when and where they received their education, any places they had traveled to for more than three months, scores on the National Test of Oral Putonghua Proficiency (if applicable), other languages they had learned, and self-estimates of Southern Min and Mandarin proficiency (1-10 scale). As for the amount of Mandarin use, the participants were requested to report the percentage of use with their grandparents, parents, children, other family members and relatives, teachers and classmates, coworkers, friends, use at home and outside home, and to estimate their overall use of Mandarin. The overall information of LBQ by age group is summarized in Table 1.

Table 1. Language background of the three age groups of Quanzhou Southern Min (SM)-Mandarin (MD) bilinguals. Mean and standard deviation (in parentheses) are shown for the age at recording, years of education, age of learning Mandarin, amount of Mandarin use, self-estimated Southern Min proficiency and self-estimated Mandarin proficiency.

Age Group	Age at recording	Edu. years	MD AOL	MD use	SM self-estimate	MD self-estimate
Younger	19.9	14	4.0	63%	7.6	8.5
	(1.1)	(0.9)	(0.8)	(10%)	(0.9)	(0.8)
Mid-age	39.6	16.9	6.3	43%	9.1	8.1
	(3.4)	(2.5)	(1.0)	(13%)	(0.8)	(1.1)
Older	58.6	11.5	7.8	25%	9.8	7.5
	(3.2)	(2.9)	(0.5)	(15%)	(0.5)	(1.3)

Table 1 indicates that the younger group learned Mandarin earlier and used it more than the mid-age group and the mid-age group learned Mandarin earlier and used it more than the older group. Both the older and mid-age groups estimated their proficiency in Southern Min to be higher than their proficiency in Mandarin, but the younger group reported higher proficiency in Mandarin than in Southern Min. All the groups reported speaking relatively more Southern Min at home and relatively more Mandarin outside the home (See Appendix B). Nevertheless, younger speakers reported more Mandarin use with parents, relatives, teachers, classmates, coworkers, and friends. The mid-age group

reported speaking more Mandarin to their children than the older group, which also implied more Mandarin use in the younger group.

2.2 Stimuli

Materials were adapted from Xu et al. [2012]. Participants were instructed to say a target sentence in both Mandarin and Southern Min, as shown in Table 2. In both languages, the target sentence had three words and five syllables. Previous studies [Xu, 1999; Xu et al., 2012] found that the presence or absence of PFC is most clearly seen in a sentence consisting of syllables with only sonorant onsets and carrying mostly level tones, like the ones used here.

Table 2. Target sentences in Mandarin and Quanzhou Southern Min. The numbers represent surface tone values based on a 5-point tone scale [Chao, 1930]

	Word 1	Word 2	Word 3
Character	妈妈	摸	猫咪
Gloss	'mom'	'pat'	'kitty'
Mandarin	ma55 ma3	mo55	mau55 mi55
QZ S. Min	ma22 ma24	m□33	niau33 mi24

All syllables of the target sentence in Mandarin had the high-level tone 55 underlyingly, but the second syllable /ma/ was realized with a neutral tone due to a rule of reduplication, so that the surface pitch level was 3 [Chen and Xu, 2006]. In Quanzhou Southern Min, /ma/ had the rising tone 24 underlyingly, but the first syllable /ma/ was realized with the surface low-level tone 22 due to a tone sandhi rule [Lin, 1993: 60]. Both /m□/ and /niau/ had the mid-level tone 33 and /mi/ had the rising tone 24 in both underlying and surface forms.

A picture illustrating the target sentence ('Mom is patting kitty') was shown to the participants in order to set up a focus-eliciting situation. Four prompt questions were used to elicit four types of focus: no focus, initial focus (on Word 1), medial focus (on Word 2) and final focus (on Word 3). In each trial, the participant read aloud the target sentence in Table 2 as a response to one of the prompt questions in Table 3. As found previously [Chen and Xu, 2006; Cooper et al., 1985; Jin, 1995; Wang et al., 2011], these questions would automatically trigger the types of focus indicated in Table 3.

Table 3. Prompt questions for eliciting foci in both languages.

Focus	Prompt questions	English translation
None	图中你看到什么?	What do you see in the picture?
Initial	谁摸猫咪?	Who is patting the kitty?
Medial	妈妈对猫咪做什么?	What is Mom doing to the kitty?
Final	妈妈摸什么?	What is Mom patting?

2. 3 Recording

The prompt questions were asked by the experimenter (first author, who is bilingual in Southern Min and Mandarin) in the relevant target language (Southern Min or Mandarin) and each question was repeated five times in a random order. The experimenter instructed the subjects to use the target sentence to answer the prompt questions as naturally as possible. The inter-trial interval (interval between adjacent question-answer pairs) was about three seconds. Participants answered the questions with the target sentence with appropriate focus. Each participant produced 40 sentences (4 foci × 5 repetitions × 2 languages). The experimental instructions were given in both Mandarin and Southern Min. The Mandarin production was recorded prior to Southern Min, except for two speakers in the older group, who preferred using Southern Min first. The recording in each language lasted about five minutes and there was a two-minute break between the two recordings. Recording was conducted in a quiet room with a Marantz professional solid state recorder PMD660 and a Shure professional unidirectional head-worn dynamic microphone. The utterances were directly recorded into a computer SD card with a sampling rate of 44,100Hz.

2.4 Analyses

Data were analyzed with Praat (version 5.1.32). ProsodyPro — a Praat script for large-scale prosody analysis [Xu, 2013] — was used to extract continuous F_0 contours as well as various measurements from each utterance. Measurements used in the present study comprised maximum F_0, minimum F_0, mean F_0, intensity, duration, and time-normalized F_0 with 10 points in each syllable interval. The time-normalized F_0 contours were used only in the graphical analysis, and all the other F_0 measurements were taken from the original non-time-normalized F_0 contours. To assess the effect of focus, differentials in mean F_0, intensity and duration between in-focus, pre-focus and post-focus words and their unfocused counterparts were calculated. The in-focus differential was calculated as the mean of measured values of the focused syllables minus that of their counterparts in the unfocused utterances. The pre-focus change and post-focus

change were calculated as the mean differences between the pre-focus or post-focus syllables and their counterparts in the unfocused utterances. For instance, post-focus differential in Mandarin was the grand mean of the values of *mo55*, *mau55* and *mi55* in the initial-focus condition (i.e., when *ma55 ma3* is focused) minus the values of the three syllables in the unfocused sentence, and the values of *mau55* and *mi55* in the medial-focus condition (i.e. when *mo55* is focused) minus the values of the two syllables in the unfocused sentence (see Table 2). These measurements therefore enabled comparisons of F_0, intensity and duration patterns across the sentences in the unfocused, initial-focus, medial-focus and final-focus conditions.

Note that, of the various acoustic measurements used in different studies, the abovementioned pre-focus/post-focus changes based on cross-utterance comparisons are among the most conservative. This is because, any difference across the focus conditions has to be larger and more consistent than both cross-utterance and cross-speaker variations for it to reach statistical significance. In contrast, within-utterance differences, though easier to observe, are prone to confounding with non-focus-related factors such as lexical stress, downstep, declination, phrasing, sentence type, etc. On the other hand, despite being conservative, PFC measured this way has been consistently found in Mandarin [Xu, 1999], English [Xu & Xu, 2005] and a number of other languages as summarized in Xu et al. [2012], indicating the robustness of PFC in these languages. By using the same conservative measurements in the present study, we intend to make sure that the presence/absence of PFC we report is based on a most stringent set of criteria.

3. Results

3.1 F_0 contours

Before making statistical comparisons, time-normalized F_0 contours of the stimulus sentences were first examined for an assessment of the overall differences across the experimental conditions. Figures 1 and 2 show the mean time-normalized contours associated with different focus conditions for the different age groups and languages.

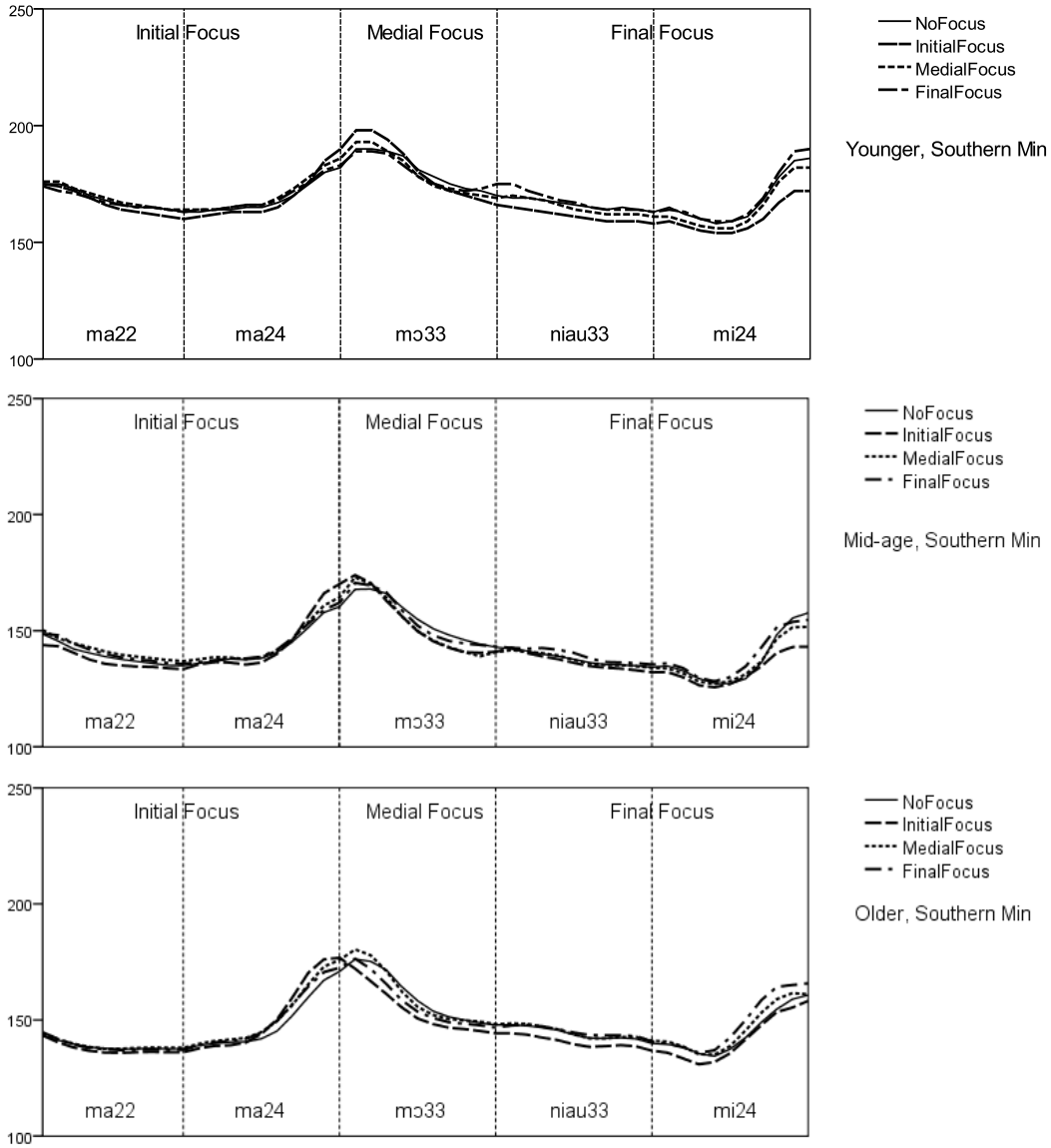

Fig.1. Mean time-normalized F0 contours in Southern Min by three age groups. Each curve represents an average of the five repetitions by eight speakers under the same focus condition. Syllable boundaries are marked with vertical dashed lines.

Figure 1 indicates that none of the three age groups produced either noticeable F_0 expansion on in-focus constituents or F_0 compression on post-focus constituents under any of the focus conditions in Southern Min.

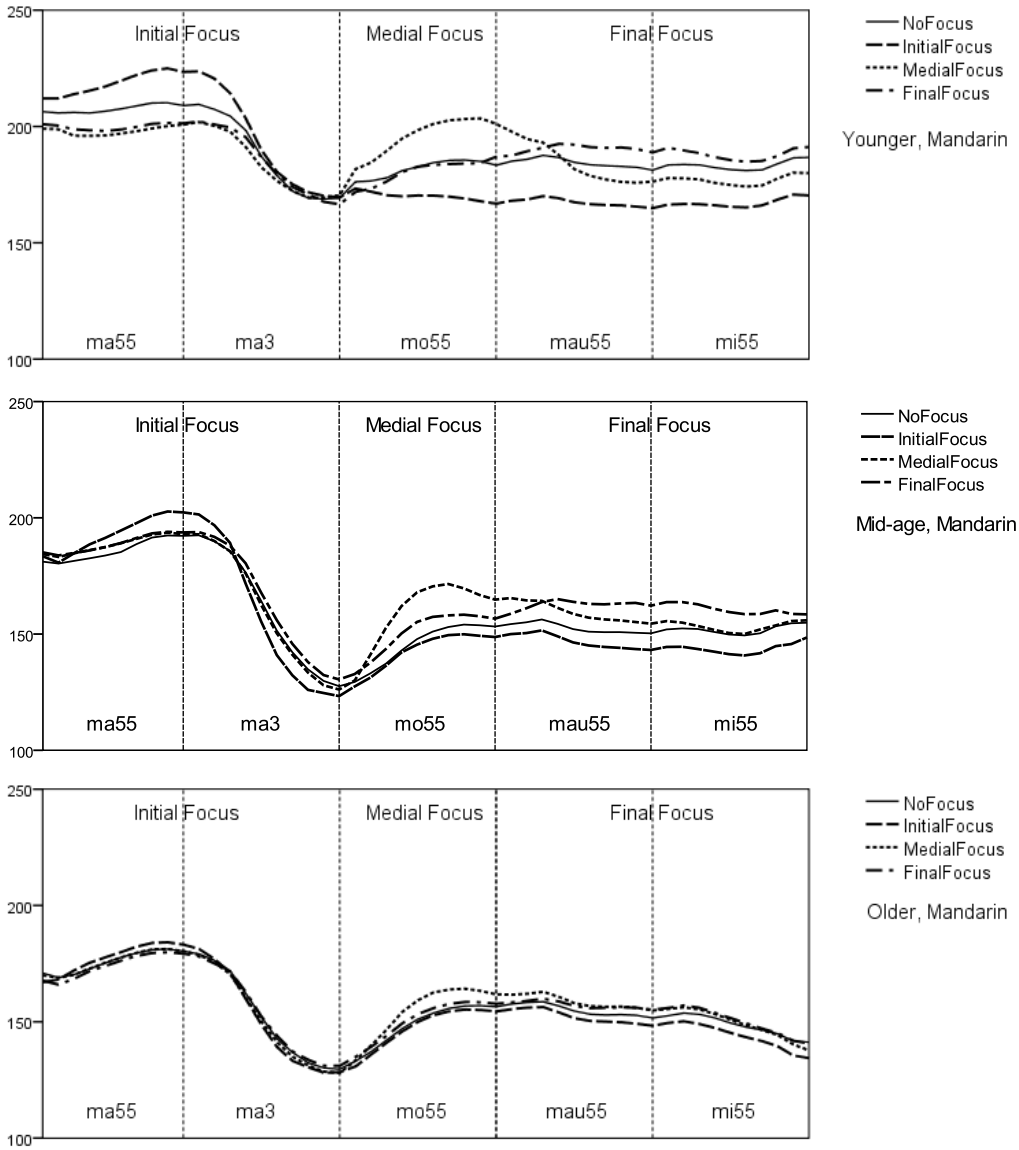

Fig. 2. Time-normalized F_0 contours in Mandarin by the three age groups. Each curve represents an average of the five repetitions by the eight speakers under the same focus condition. Syllable boundaries are marked with vertical dashed lines.

Figure 2 indicates that in-focus words produced by the younger and mid-age groups had higher F_0 than their unfocused counterparts in Mandarin. In the younger group, post-focus F_0 contours in sentences with both initial focus and medial focus were lower than their counterparts in the unfocused utterances. In the mid-age group, post-focus F_0 is lower than unfocused F_0 in sentences with initial focus but not in sentences with medial focus. The older group did not show clear post-focus lowering of F_0 or any other noticeable F_0 differentials relative to the unfocused utterances.

3.2 Differentials in mean F_0, intensity, and duration

To statistically verify the F_0 differentials over four focus conditions across the three age groups in both languages as seen in Figures 1 and 2, mean F_0 differences were converted from Hz to semitones ($st = 12\ log2(F_0)$, where reference level is 1 Hz) and compared in repeated measures ANOVAs. Conversion to semitones is necessary because pitch in speech operates on a logarithmic scale just as in music [Fujisaki, 2003; Nolan, 2003]. The dependent variable is F_0 change from the unfocused condition to the focus condition. The three independent variables are the between-subjects factor — age group (younger, mid-age, older) — and the within-subjects factors—language (Southern Min, Mandarin) and focus condition (pre-focus, in-focus, post-focus). Individual speaker is the random factor in all these ANOVAs. The results showed a three-way interaction between the factors on mean F_0 differential ($F(4,42) = 4.436$, $p = 0.004$) and two-way interactions between language and focus condition ($F(2,42) = 17.233$, $p < 0.001$) and between focus condition and age ($F(4,42) = 2.694$, $p = 0.044$). Figure 3 displays the means and standard errors according to focus condition and age group in the two languages.

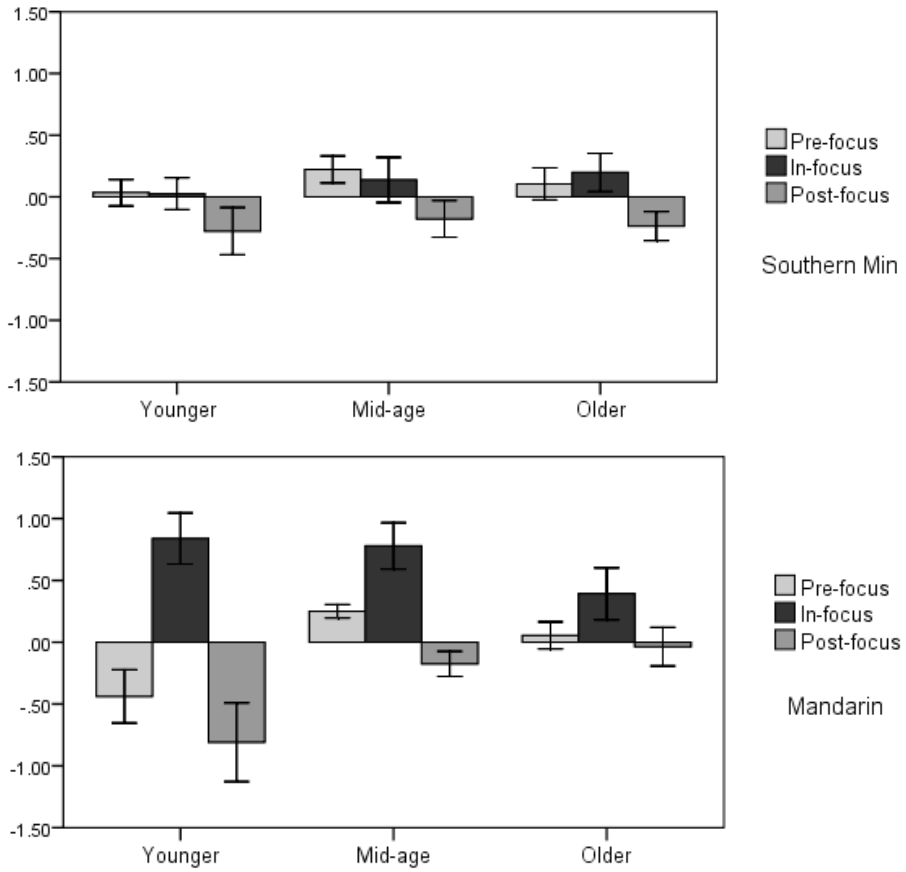

Fig. 3. Mean F_0 differential (semitone) by focus condition and age group in both languages.

The data were then split by language to further examine the effect of age group and focus condition within each language. The F_0 data in Southern Min showed no interaction between age group and focus condition and no significant main effect of age group on F_0 differential. However, the main effect of focus condition on F_0 differential was significant ($F(2,21) = 15.679$, $p < 0.001$). The source of the effect of focus condition on F_0 differential seems to be some slight variations of pre-focus and especially post-focus F_0 differential compared to small in-focus F_0 differentials by all the three groups (upper panel in Figure 3). To test the effect of focus condition, the magnitude of F_0 differential was calculated by subtracting the in-focus differential respectively from the pre-focus differential and from the post-focus differential. However, independent samples t-tests showed no significant difference of either magnitude of F_0 differential between any two age groups, suggesting that none of the age groups used F_0 differential to code focus in Southern Min.

The Mandarin data on F_0 differential showed significant interaction between age group and focus condition ($F(4,42) = 4.518$, $p = 0.004$) and significant main effects of both focus condition ($F(2, 21) = 34.685$, $p < 0.001$) and age group ($F(2, 21) = 3.625$, $p = 0.044$). Post-hoc independent samples t-tests were run for group comparison in each focus differential. Since there are three comparisons among the three groups with a familywise error rate of .05, we ran each test at the .05/3 = .0167 level.[2] The results showed significant differences in pre-focus F_0 differential between younger and mid-age groups ($t(14) = -3.097$, $p = 0.015$) and between younger and older groups ($t(14) = -2.798$, $p = 0.014$). Also a near-significant difference in post-focus F_0 differential was found between younger and older groups ($t(14) = -2.719$, $p = 0.022$). The younger and older groups also demonstrated significant differences of the magnitude of F_0 differential between pre-focus and in-focus positions ($t(14) = -2.971$, $p = 0.010$) and between post-focus and in-focus positions ($t(14) = -3.351$, $p = 0.005$). The lower panel in Figure 3 indicates that the younger group produced both pre-focus and post-focus compression of F_0; the mid-age group did not significantly reduce post-focus F_0 and even increased pre-focus F_0; the older group did not have any significant variation of F_0 on pre-focus and post-focus words and even did not expand F_0 as much as the younger and mid-age groups. The overall F_0 pattern of the younger group in Mandarin was therefore different from those of mid-age and older groups.

Repeated measures ANOVAs similar to those on F_0 differential were also conducted on mean intensity differential. No three-way interaction was found among the three factors: language, age group and focus condition. There were interactions between language and age group ($F(2,21) = 5.909$, $p = 0.009$) and between age group and focus condition ($F(4,42) = 5.080$, $p = 0.002$). Figure 4 displays the pattern of intensity differential for the two languages divided by focus condition and age group.

[2] This kind of downward adjustment of p values for multiple comparisons has been argued to be overly conservative [Perneger, 1998].

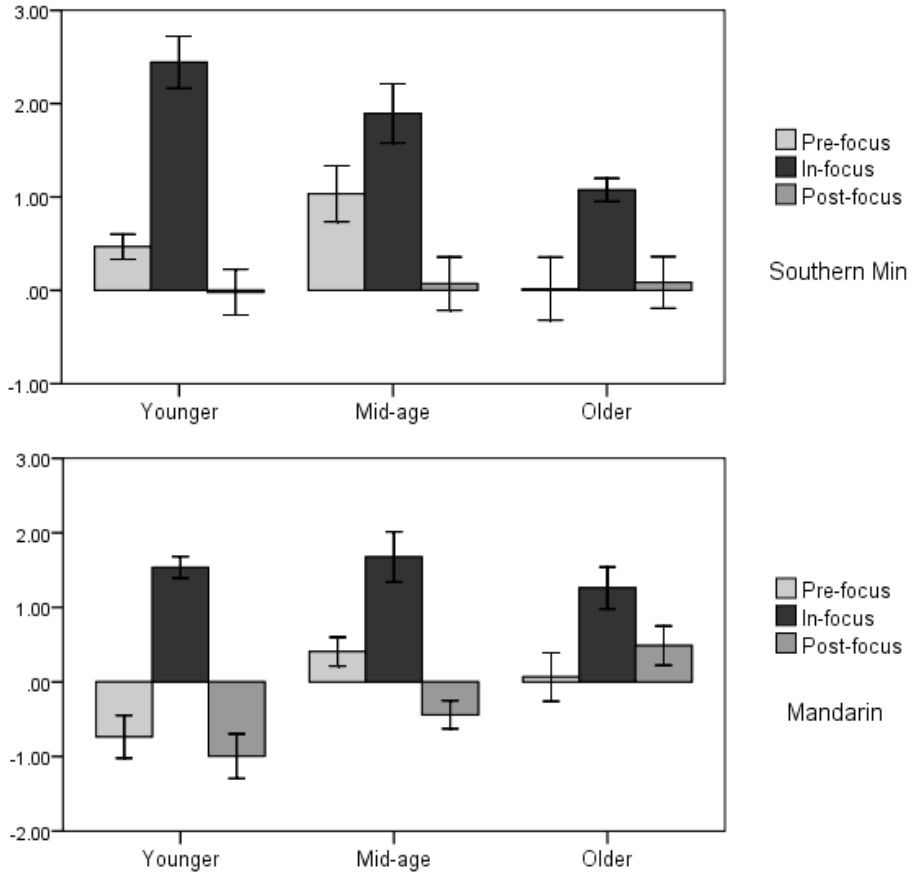

Fig. 4. Mean intensity differential (dB) by focus condition and age group in both languages.

The Southern Min data on intensity differential showed a two-way interaction between age group and focus condition ($F(4,42) = 3.466$, $p = 0.016$) and significant main effects of focus condition ($F(2, 21) = 44.507$, $p < 0.001$) and age group ($F(2, 21) = 3.710$, $p = 0.042$). Post-hoc independent samples t-tests showed a significant difference in in-focus intensity differential between younger and older groups ($t(14) = 5.502$, $p < 0.001$). The magnitude of intensity differential between post-focus and in-focus positions was significantly different between the younger and older groups ($t(14) = -4.707$, $p < 0.001$). The younger and mid-age groups demonstrated a near significant difference in the magnitude of intensity differential between pre-focus and in-focus positions ($t(14) = -2.573$, $p = 0.022$). The upper panel of Figure 4 shows that the younger group increased intensity in the in-focus position more than mid-age and older groups and the mid-age group increase intensity in the pre-focus position more than younger and older groups. However, none of the groups showed post-focus compression of intensity in Southern Min.

The Mandarin data on intensity differential shows a two-way interaction between focus condition and age group ($F(4,42) = 3.995$, $p = 0.008$) and significant main effects of focus condition ($F(2, 21) = 43.292$, $p < 0.001$) and age group ($F(2, 21) = 5.437$, $p =$

0.013). Post-hoc independent samples t-tests showed significant differences in pre-focus intensity differential between younger and mid-age groups ($t(14) = -3.308$, $p = 0.005$) and post-focus intensity differential between younger and older groups ($t(14) = -3.728$, $p < 0.001$) and between mid-age and older-group ($t(14) = -2.874$, $p = 0.012$). Intensity differential between post-focus and in-focus positions was significantly different between younger and older groups ($t(14) = -4.222$, $p = 0.001$) and between mid-age and older groups ($t(14) = -3.583$, $p = 0.003$). The lower panel in Figure 4 shows that in Mandarin all the three groups increased intensity on in-focus words, but only the younger and mid-age groups reduced post-focus intensity, and the younger group also reduced pre-focus intensity.

A final set of repeated measures ANOVAs similar to those on F_0 and intensity differentials conducted on duration differential showed no three-way interaction and no two-way interactions between any of the factors. The main effects of age group on duration differential were not significant either. However, there were significant main effects of language ($F(1,21) = 4.507$, $p = 0.046$) and focus condition ($F(2,42) = 47.896$, $p < 0.001$). Figure 5 shows a large increase in duration on focused words and small differentials in duration on pre-focus and post-focus words in both Southern Min and Mandarin.

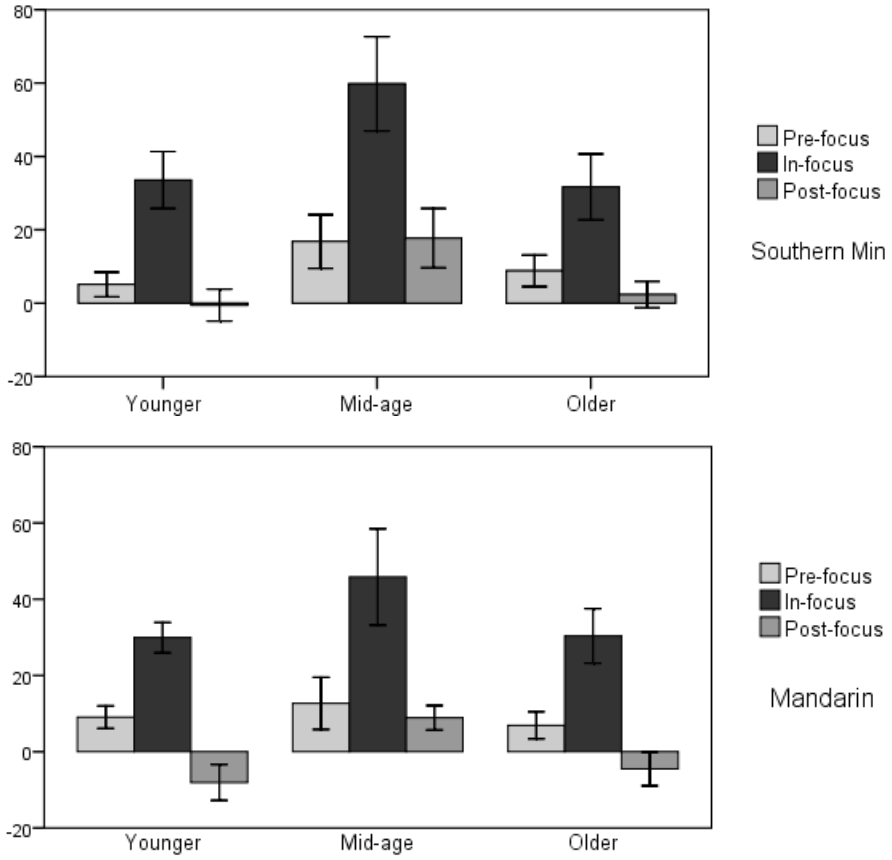

Fig. 5. Mean duration differential (ms) by focus condition and age group in both

languages.

The Southern Min data showed no interaction between age group and focus condition and no significant main effect of age group on duration differential. However, the main effect of focus condition on duration differential was significant ($F(2,21) = 46.984$, $p < 0.001$). Furthermore, duration differential neither between pre-focus and in-focus positions nor between post-focus and in-focus positions was significantly different between any two groups in Southern Min.

As in Southern Min, the duration data in Mandarin showed no interaction between age group and focus condition nor significant main effect of age group on duration differential. But there was a significant main effect of focus condition on duration differential ($F(2,21) = 35.166$, $p < 0.001$). Duration differentials between pre-focus and in-focus positions and between post-focus and in-focus positions were not significantly different between any two groups in Mandarin.

3.3 Correlations between language experience and PFC for F_0 and intensity

As expected, more experience in L2 — specifically, greater L2 Mandarin use — resulted in the production of PFC in Mandarin: no older speakers produced significant PFC and more younger speakers produced PFC than mid-age speakers. Both Figures 3 and 4 show, however, that there is much cross-speaker variability within each group. To examine how much of this variability is related to L2 use, a final set of analyses was conducted. Linear regressions were performed with language use reported by individual speakers as the predictor, and PFC in Mandarin as the dependent variable, as shown in Figures 6 and 7. Since age of learning (AOL) L2 also varied somewhat between the groups, it was used as a control variable. Here PFC was calculated by subtracting the unfocused mean F_0 and intensity from the post-focus means.

L2 use significantly predicted PFC of F_0 ($\beta = -0.657$, $p = 0.007$) and also explained a significant proportion of variance in PFC of F_0 ($R^2 = 0.482$, $F(2, 21) = 9.827$, $p = 0.001$). L2 use was also a significant predictor of PFC of intensity ($\beta = -0.746$, $p = 0.001$) and explained a significant proportion of variance in PFC of intensity ($R^2 = 0.576$, $F(2, 21) = 14.289$, $p < 0.001$). In contrast, AOL was not a significant covariant of PFC of either F_0 or intensity. The regression plots in Figures 6 and 7 show that the amount of L2 use varied across the three age groups, although there are also clear overlaps. In both Figures 6 and 7, there seems to be an even stronger relationship between Mandarin use and PFC in the younger speakers. This might suggest that only when L2 proficiency exceeds some threshold does PFC start to emerge. But none of the nonlinear functions we tested generated better predictions than the linear functions shown here.

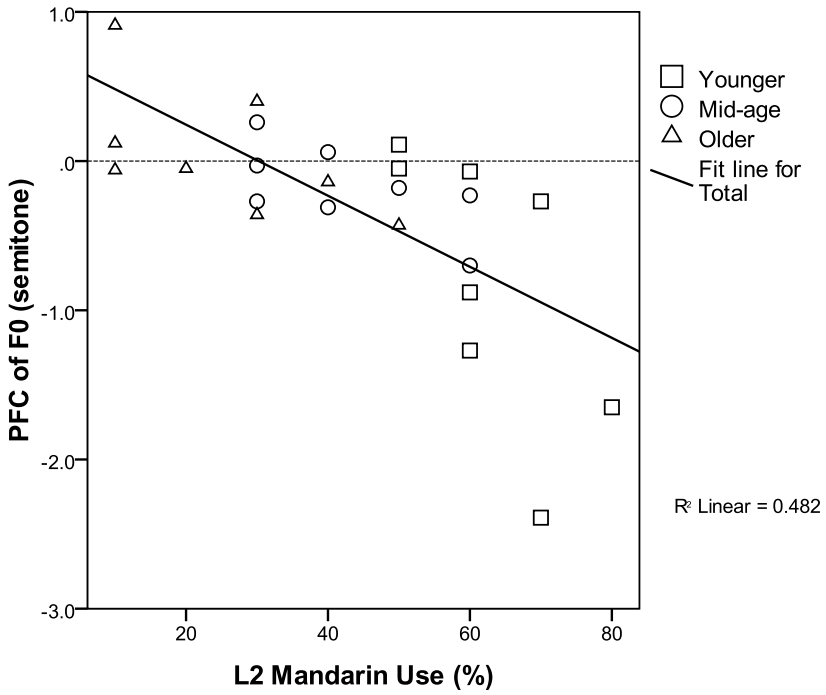

Fig. 6. Regression plots of PFC of F_0 in Mandarin as a function of L2 Mandarin use for all subjects.

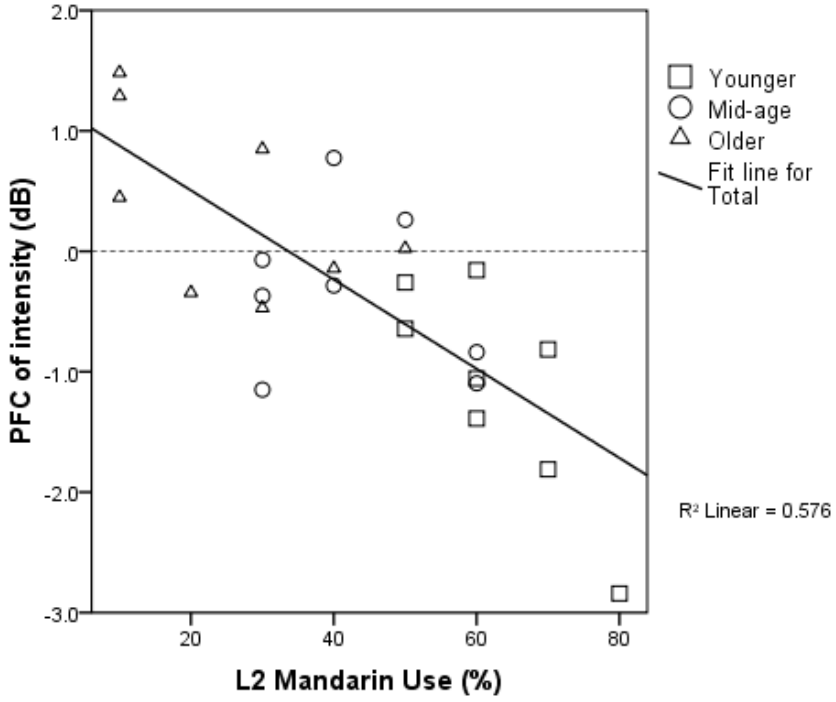

Fig. 7. Regression plots of PFC of intensity in Mandarin as a function of L2 Mandarin use for all subjects.

Finally there is also a highly significant correlation between PFC of F_0 and PFC of intensity ($r(24) = 0.756$, $p < 0.001$). This is consistent with the findings of Xu et al. [2012], which suggests that PFC of F_0 and PFC of intensity may not be independent of each another.

4. Discussion

The results just presented have provided answers to the four questions investigated in this study. First, all age groups expanded duration and intensity on the focused words in both Southern Min and Mandarin. This is consistent with the findings of Pan [2007] and Xu et al. [2012] for Taiwan Southern Min and those of Jin [1996] and Xu [1999] for Mandarin. However, unlike in Mandarin, none of the Quanzhou bilinguals expanded F_0 for in-focus items in Southern Min. This result differs from the finding in a conversational corpus study by Pan [2007] that Taiwan Southern Min increase F_0 on the in-focus item. But it is at least in line with her conclusion that in-focus F_0 increase is not as consistent as duration lengthening in Southern Min.

Second, PFC was not found in Southern Min for any of the age groups of Quanzhou speakers in the current study. Although all Quanzhou bilingual groups expanded duration, intensity and F_0 on focused words in Mandarin, only the younger group produced significant PFC in F_0 and intensity.

Third, as shown in Figures 2, 3 and 4 and in the statistical results, in contrast to the older group, which produced no consistent PFC in their Mandarin, the younger group showed clear PFC. And, just as interestingly, the mid-age group showed a greater amount of PFC than the older group. The mid-age group may therefore represent a transitional or intermediate stage of realizing prosodic focus that lies somewhere between that of Beijing Mandarin and Southern Min and between the production of younger and older generations. The results also show that younger Quanzhou bilinguals speaking Mandarin tend to compress F_0 and intensity in the pre-focus constituents. This pre-focus compression goes beyond the typical Mandarin norms and may reflect an extra effort in this group to realize focus as clearly as possible.

The differences by age group are likely to be attributable to differences in language experience, specifically, the amount of L2 use (see Figures 6 and 7), since all groups were exposed to Mandarin at similar ages and age of learning was found not to be a good predictor of PFC. The three groups also share many other similarities, being all from similar societal communities, some even from the same families. These similarities include style of upbringing, inter-personal interaction, education, economic situation, political environment, mass media and contact with broader circles of places and people. Data from the language background questionnaire (LBQ) indicated that age of learning Mandarin differed by only about two years between adjacent age groups and all the speakers started learning Mandarin in their early childhood. All speakers can therefore be considered “early learners” [Flege, MacKay and Meador, 1999; Guion, 2003; Hojen and

Flege, 2006]. Data from the LBQ indicated that the amount of L1/L2 use differed by about 20% between the age groups – the younger the speakers, the more the Mandarin use (See Appendixes A and B). This difference appears to have influenced speakers' self-assessments of language proficiency. Younger speakers indicated higher proficiency in Mandarin than mid-age and older speakers. Younger speakers also thought their Mandarin proficiency was higher than their Southern Min proficiency. Finally, the finding that self-estimated L2 use predicted a significant proportion of the variance in post-focus compression of F_0 and intensity in Mandarin supports the hypothesis that L2 use is the primary factor responsible for the acquisition of Mandarin focus pattern by native speakers of Southern Min. The prediction of PFC by L2 use was far from perfect, however, suggesting that there are also other contributing factors that are out of the scope of the present study. Identification of these additional factors will require future studies designed to examine them specifically.

The above-mentioned results can now allow us to address the predictions related to research questions 2 and 4 outlined in the Introduction. Regarding predictions 1a and 1b, the results are consistent with the latter, i.e., *PFC occurs in L2 Mandarin when there is sufficient L2 experience*. Some of the Quanzhou bilinguals in the present study, especially those in the younger group, produced PFC in their L2 Mandarin. This is different from the finding of Xu et al. [2012] that none of the speaker groups in Taiwan showed PFC in their Mandarin. Here an important factor could be the type of L2 input. Flege and Liu [2001] found that even adults' L2 performance could be improved if they received a substantial amount of input from native speakers. However, the amount of formal instruction in the L2 was not a positive factor for L2 pronunciation accuracy [Piske, MacKay and Flege, 2001; Piske, 2007]. Those findings suggest that a large amount of L2 input is not sufficient to achieve a native-like accent if the speakers were exposed to non-native L2 speech. The younger speakers in the present study were immersed more in the Mandarin-speaking environment than mid-age or older speakers. They also received more Beijing-Mandarin-like input, due to the surge in Mandarin media during the younger speakers' lifetime, than mid-age and older speakers. The older speakers, on the other hand, as the first generation of Southern Min-Mandarin bilinguals in Mainland China, did not receive Beijing-Mandarin-like input as the younger generation did, for the teachers of the older generation probably had a lower Beijing Mandarin proficiency than the younger generation's teachers. In addition to receiving better input, most of the speakers in the younger age group and some speakers in the mid-age group were required to take the National Test of Oral Putonghua Proficiency for their future or current profession. These speakers also received intensive training in class before they took the associated oral test. Intensive training at both segmental and suprasegmental levels has been found to be effective in decreasing foreign accent [Moyer, 1999; Missaglia, 1999; Piske et al., 2001]. The effect of special training may have supplemented the effect of higher quality input to ensure that younger speakers were Beijing-like in their L2

productions of focus in Mandarin. Relating these findings to the total lack of PFC in Taiwan Mandarin regardless of language experience [Xu et al., 2012], it is likely that the Mandarin L2 input in Taiwan is different from that in Quanzhou. In that study, even the monolingual Mandarin speakers showed no sign of PFC, which suggests that there is a widespread loss of this prosodic pattern in the Taiwan version of Mandarin. Further research is needed to look into this issue.

Regarding predictions 2a and 2b, the results of the present study support 2b, namely, *PFC does not occur in L1 Southern Min, regardless of L2 Mandarin experience*. There was no effect of L2 Mandarin prosody on the production of L1 Southern Min prosody production in any of the three age groups. This differs from previous findings that, at the segmental level, there are mutual effects of L1 and L2 sound systems in societal bilinguals in vowel [Guion, 2003] and consonant [Peng, 1993) production. The lack of L2 to L1 impact in the present study is especially notable for the younger speakers, whose Mandarin use was as high as 63% on average (Table 1) and who showed clear PFC in their Mandarin. This is consistent with the hypothesis that post-focus compression is not easily transferred from one language to another [Wu and Chung, 2011; Xu, 2011]. But the findings of the present study are by no means the final answer to the question of transferability. For example, what if L2 use is even more than what is observed in the present study? And what happens when speakers whose L1 has PFC try to learn a non-PFC L2? More studies are needed along this new line of inquiry.

Finally, a note on methodology is in order, as there are often questions about the validity of investigating prosodic focus in the laboratory as done in the present study. One concern is that it is unnatural to say the same sentence repeatedly in the laboratory, and it would be much better to examine focus in naturally occurring speech. Regarding this, we first need to realize that if the participants have received only procedural instructions, as is the case in the present study, the prosodic patterns they displayed could only reflect their own natural way of realizing focus. As for observing naturalistic speech directly, the question is how to collect data that can be subjected to the same level of scrutiny as applied in the present study. That is, first, the location and scope of focus in each natural utterance needs to be as certain as those in the present study, where each utterance was produced after a prompt question. Second, the utterances being compared need to be similar to each other in terms of context, semantics, utterance length, syntactic structure, etc., so as to be free of confounding effects from these factors. Third, utterances produced by different speakers need to be sufficiently similar to each other for any conclusion about speaker group to be valid. It is not hard to imagine how difficult it would be to collect a natural speech corpus that satisfied all these requirements.

If laboratory data collection is inevitable, however, a further concern is how we can know that PFC, or any prosodic pattern found in a laboratory study, would match natural speech. For this we have to admit that one can never achieve a hundred percent certainty, because the difficulty is again how to obtain naturalistic data that are directly comparable

to the laboratory data. A possible alternative is to use computational modeling that can simulate the learning and production of prosody directly from naturalistic data, using machine learning tools that have been developed recently [Prom-on, Xu and Thipakorn, 2009; Xu and Prom-on, 2014]. This again awaits future studies.

5. Conclusions

The current study examined prosodic realization of focus in the production of Southern Min-Mandarin bilinguals in Mainland China. Unlike the Taiwan Mandarin speakers in Xu et al. [2012], the younger Quanzhou bilinguals in this study produced more Beijing-like prosodic focus in Mandarin than older speakers, i.e., with substantial PFC in F_0 and intensity. As in Xu et al. [2012], however, all age groups produced Taiwan-like prosodic focus in Southern Min, i.e., with no significant PFC of F_0 or intensity. Meanwhile, all the age groups presented comparable patterns of duration in their L2 Mandarin as well as L1 Southern Min production.

Despite the variability in the background of the three age groups, such as family, education, economy, work environment, etc., the general situation of language experience in Quanzhou bilinguals is that the younger generation use more L2 Mandarin than the older generation and the amount of Mandarin use does not significantly vary within generation. Therefore, the results of the current study suggest language experience affects prosodic realization of focus in Southern Min-Mandarin bilinguals in mainland China. The current study also suggests that PFC can be learned given sufficient language experience, i.e., early exposure to high-quality L2 input and extensive use of L2. There could also be other factors that may accelerate PFC learning even further, but identifying them will require future studies with appropriate designs. The present findings also suggest that PFC could be used as an index in the assessment of second language speech proficiency, especially in the case of speakers of non-PFC language learning a PFC language. Previously, L2 efficiency has been mostly examined by subjective ratings of global pronunciation by native speakers [Flege, Frieda and Nozawa, 1997; Guion, Flege and Loftin, 2000; Piske, MacKay and Flege, 2001]. The current study examined the effect of L1/L2 using detailed acoustic measurements instead of subjective ratings. It would be interesting, however, to further examine how well acoustically measured PFC matches L2 proficiency judged by native listeners.

Acknowledgements

Our thanks go to Melissa Redford, Melissa Baese-Berk, Vsevolod Kapatsinski, Koari Idemaru, Jonathan Loftin and the anonymous reviewers for their valuable comments on an earlier version of this paper. We also thank the audience at the 160th

meeting of the Acoustical Society of America and the 6th International Conference of Speech Prosody for their comments on part of the data in this work.

References

Baker, W.; Trofimovich, P: Interaction of native- and second-language vowel system(s) in early and late bilinguals. Lang. Speech *48*: 1-27 (2005).

Bolinger, D.: Accent is predictable (if you're a mind-reader). Language *37*: 83-96 (1972).

Chao, Y-R.: A system of tone-letters. Le maître phonétique *45*: 24–27.

Chen, Y.; Xu, Y.: Production of weak elements in speech -- Evidence from F_0 patterns of neutral tone in standard Chinese. Phonetica *63*: 47-75 (2006).

Cooper, W. E.; Eady, S. J.; Mueller, P. R.: Acoustical aspects of contrastive stress in question-answer contexts. J. acoust. Soc. Am. *77*: 2142-2156 (1985).

Eady, S. J.; Cooper, W. E.: Speech intonation and focus location in matched statements and questions. J. acoust. Soc. Am. *80*: 402-415 (1986).

Féry, C.: Focus as prosodic alignment. Nat. Lang. & Ling. Theo. *31*: 683-734 (2013).

Flege, J. E.: Second language speech learning: Theory, findings, and problems, in W. Strange, Speech perception and linguistic experience: Issues in cross-language research, pp. 233-277 (York Press, Timonium, MD 1995).

Flege, J. E. (2007). Language contact in bilingualism: Phonetic system interactions; in J. Cole & J. Hualde, Laboratory Phonology 9, pp. 353-382 (Berlin: Mouton de Gruyter).

Flege, J. E.; Frieda, E. M.; Nozawa, T.: Amount of native language (L1) use affects the pronunciation of an L2. J. Phonet. *25*: 169-186 (1997).

Flege, J. E.; MacKay, I. R. A.; Meador, D.: Native Italian speakers' production and perception of English vowels. J. acoust. Soc. Am. *106*: 2973-2987 (1999).

Flege, J. E.; Munro, M. J.; MacKay, I. R. A.: Factors affecting degree of perceived foreign accent in a second language. J. acoust. Soc. Am. *97*: 3125-3134 (1995).

Fujisaki, H.: Prosody, Information, and Modeling — with Emphasis on Tonal Features of Speech; Workshop on Spoken Language Processing, pp. 5-14. (2003).

Guion, S. G.: The vowel systems of Quichua-Spanish bilinguals: An investigation into age of acquisition effects on the mutual influence of the first and second languages. Phonetica *60*: 98-128 (2003).

Guion, S. G.; Flege, J. E.; Loftin, J.: The effect of L1 use on pronunciation in Quichua-Spanish bilinguals. J. Phonet. *28*: 27-42 (2000).

Guion, S. G.; Flege, J. E.; Liu, S.; Yeni-Komshian, G: Age of learning effects on the

duration of sentences produced in a second language. Appl. Psycholing. *21*: 205-228 (2000).

Gussenhoven, C.: Focus, mode and the nucleus. J. Ling. *19*: 377-417 (1983).

Hanssen, J.; Peters, J.; Gussenhoven, C.: Prosodic effects of focus in Dutch declaratives. ISCA Arch., Speech Pros. (Campinas, Brazil 2008).

He, X.; Hanssen, J.; van Heuven V. J.; Gussenhoven, C.: Phonetic implementation must be learnt: native versus Chinese realization of focus accent in Dutch. Proc. 17th ICPhS 843-846 (Hong Kong 2011).

Hojen, A.; Flege, J.: Early learners' discrimination of second-language (L2) vowels. J. acoust. Soc. Am. *119*: 3072-3084 (2006).

Huang, B. H.; Jun, S-A.: The effect of age on the acquisition of second language prosody. Lang. Speech *54* (3), 387-414 (2011).

Huang, Y-H.; Fon, J.: The effect of acquisition order and word relatedness on code-switching costs in balanced bilingual speakers. Proc.16th ICPhS: 1577-1580 (Saarbrücken, Germany 2007).

Jin, S.: An acoustic study of sentence stress in Mandarin Chinese. PhD diss. The Ohio State University (1996).

Kang, K.-H.; Guion, S.: Phonological systems in bilinguals: Age of learning effects on the stop consonant systems of Korean-English bilinguals. J. acoust. Soc. Am. *119*: 1672-1683 (2006).

Kubler, C. C.: The influence of Southern Min on the Mandarin of Taiwan. Anthrop. Ling. *27.2*: 156-176 (1995).

Kubozono, H.: Focus and intonation in Japanese: Does focus trigger pitch reset?; in Ishihara, S., Jannedy, S. Schwarz, A., Working Papers of the SFB632, ISIS *9*: 1-27 (Universitätsverlag Potsdam 2007).

Ladd, D. R.: Intonational phonology. (CUP, Cambridge 1996).

Lambrecht, K.: Information structure and sentence form (Cambridge University Press, Cambridge 1994).

Lee, B.; Guion, S.G.; Harada, T.: Acoustic analysis of the production of unstressed English vowels by early and late Korean and Japanese bilinguals. Studies in Sec. Lang. Acq.: *28*, 487-513 (2006).

Lee, Y-c.; Xu, Y.: Phonetic realization of contrastive focus in Korean. Proc. Speech Prosody2010, 100033:1-4.

Li, D. C. S.; Lee, S.: Bilingualism in East Asia; in T. K. Bhatia, W. C. Ritchia, The Handbook of Bilingualism pp.742-779 (Blackwell, Oxford 2006).

Liao, S.: A Perceptual Dialect Study of Taiwan Mandarin: Language Attitudes in the Era of Political Battle. In M. K. M. Chan, H. Kang, Proc.NACCL-20, 1: 391-408 (OSU 2008).

Lin, L.: *Quanzhoushi Fangyanzhi* (Dialect record of Quanzhou city). (Soc. Sci. Lit. Press, Beijing 1993).

McAllister, R.; Flege, J.; Piske, T.: The influence of the L1 on the acquisition of Swedish vowel quantity by native speakers of Spanish, English and Estonian. J. Phonet. *30*: 229-258 (2002).

McGory, J. T: Acquisition of intonational prominence in English by Seoul Korean and Mandarin Chinese speakers. PhD diss. The Ohio State University (1997).

Missaglia, F.: Contrastive prosody in SLA: an empirical study with adult Italian learners of German. Proc. 14th ICPhS, 1: 551-554 (San Francisco 1999).

Moyer, A.: Ultimate attainment in L2 phonology. Studies in SLA *21*: 81-108 (1999).

Munro, M. J.; Flege, J. E.; MacKay, I. R. A.: The effects of age of second-language learning on the production of English vowels. Appl. Psycholing. *17*:313-334 (1996).

Muysken, P.: Two linguistic systems in contact: grammar, phonology and lexicon; in T. K. Bhatia, W. C. Ritchia, The handbook of bilingualism, pp. 147-168 (Blackwell, Oxford 2006).

Nava, E. & Zubizarreta, M. L.: Prosodic Transfer in L2 Speech: Evidence from Phrasal Prominence and Rhythm. Proc. Speech Pros. *2008*: 335-338 (2008).

Nolan, F.: Intonational equivalence: an experimental evaluation of pitch scales; Proc. 15th ICPhS, pp. 771-774. (Barcelona 2003).

Pan, H.: Focus and Taiwanese unchecked tones; in C. Lee, M. Gordon, D. Büring, Topic and Focus: Cross-linguistic Perspectives on Meaning and Intonation, pp. 195-213 (Springer, Dordrecht 2007).

Patil, U.; Kentner, G.; Gollrad, A.; Kügler, F.; Fery, C.; Vasishth, S.: Focus, word order and intonation in Hindi. J. South Asian Ling. *1*: 55-72 (2008).

Peng, S-H.: Cross-language influence on the production of Mandarin /f/ and /x/ and Taiwanese /h/ by native speakers of Taiwanese Amoy. Phonetica *50*: 245-260 (1993).

Perneger, T. V. What's wrong with Bonferroni adjustments. British Med. J. 316: 1236-1238 (1998).

Pierrehumbert, J.: The phonology and phonetics of English intonation. PhD diss. MIT (1980).

Pierrehumbert, J.: Prosody, intonation, and speech technology; in M. Bates, R.

Weischedel, Challenges in Natural Language Processing, pp. 257-282 (Cambridge University Press, Cambridge 1993).

Piske, T.: Implications of James E. Flege's research for the foreign language classroom; in O.-S. Bohn, M. Munro, Language experience in second language speech learning pp. 301-314 (John Benjamins, Amsterdam 2007) .

Piske, T.; MacKay, I.R.A.; Flege, J.E.: Factors affecting degree of foreign accent in an L2: a review. J. Phonet. *29*: 191-215 (2001).

Prom-on, S.; Xu, Y.; Thipakorn, B.: Modeling tone and intonation in Mandarin and English as a process of target approximation. J. acoust. Soc. Am. 125: 405-424 (2009).

Romaine, S. The bilingual and multilingual community; in T. K. Bhatia, W. C. Ritchia, The handbook of bilingualism pp. 385-405 (Blackwell, Oxford 2006).

Sankoff, G.: Linguistic outcomes of language contact; in J. Chambers, P. Trudgill, N. Schilling-Estes, The handbook of language variation and change pp. 638-668 (Blackwell, Oxford 2001).

Selkirk, E.: Sentence prosody: intonation, stress, and phrasing; in J. A. Goldsmith, The handbook of phonological theory pp. 550-569 (Blackwell, Oxford 1995).

Selkirk, E.: Bengali intonation revisited: An optimality theoretic analysis in which FOCUS stress prominence drives FOCUS phrasing; in C. Lee, M. Gordon, D. Büring, Topic and focus: a cross-linguistic perspective, pp. 217-246 (Kluwer Academic Publishers, Dordrecht 2006).

Thomason, S. G.; Kaufman, T.: Language contact, creolization and genetic linguistics (UC Press, Berkeley 1988).

Vainio, M.; Järvikivi, J.: Focus in production: Tonal shape, intensity and word order. J. acoust. Soc. Am. *121*: EL55-61 (2007).

Van Coetsem, F.: Loan phonology and the two transfer types in language contact, Publications in language sciences *27* (Foris Publications, Dordrecht 1988).

Wang, B.; Wang, L.; Qadir, T.: Prosodic encoding of focus in six languages in China. Proc.17th ICPhS, Hong Kong 2011, pp 144-147.Wu, W. L.; Chung, L.: Post-focus compression in English-Cantonese bilingual speakers. Proc. 17th ICPhS, Hong Kong 2011, pp 148-151.

Wu, W. L.; Xu, Y.: Prosodic focus in Hong Kong Cantonese without post-focus compression. Proc. Speech Prosody 2010, 100040:1-4.

Xu, Y.: Effects of tone and focus on the formation and alignment of F_0 contours. J. Phonet. *27* : 55-105 (1999).

Xu, Y.: Post-focus compression: Cross-linguistic distribution and historical origin. Proc. 17th ICPhS 152-155, Hong Kong 2011, pp152-155.

Xu, Y.: ProsodyPro — A tool for large-scale systematic prosody analysis; Tools and Resources for the Analysis of Speech Prosody (TRASP 2013), pp. 7-10. (Aix-en-Provence, France 2013).

Xu, Y.; Chen, S.-w.; Wang, B.: Prosodic focus with and without post-focus compression (PFC): A typological divide within the same language family? The Linguistic Review *29***:** 131-147 (2012).

Xu, Y.; Prom-on, S.: Toward invariant functional representations of variable surface fundamental frequency contours: Synthesizing speech melody via model-based stochastic learning. Speech Commun. *57*: 181-208 (2014).

Xu, Y.; Xu, C. X.: Phonetic realization of focus in English declarative intonation. J. Phonet. *33*: 159-197 (2005).

Appendix A. Language background of all speakers in the three age groups of QZ Southern Min (SM)-Mandarin (MD) bilinguals

	Age of test	Edu. year	MD AOL	MD use	SM self estimate	MD self estimate	PTH test*
Younger							
Y-f-1	21	15	4	60%	8	9	2A
Y-f-2	20	14	4	70%	8	8	2A
Y-f-3	18	13	3	80%	6	8	2B
Y-f-4	19	13	5	60%	9	10	
Y-m-1	20	14	4	70%	7	8	2A
Y-m-2	19	13	3	60%	7	8	2B
Y-m-3	21	15	4	50%	8	8	2A
Y-m-4	21	15	5	50%	8	9	
Mean	**19.9**	**14**	**4.0**	**63%**	**7.6**	**8.5**	
SD	**1.1**	**0.9**	**0.8**	**10%**	**0.9**	**0.8**	
Mid-age							
M-f-1	35	19	6	50%	9	10	2A
M-f-2	37	14	7	60%	9	8	
M-f-3	42	16	6	30%	10	8	
M-f-4	43	16	7	60%	9	8	
M-m-1	37	22	4	40%	8	8	2B
M-m-2	37	16	6	30%	10	8	
M-m-3	42	16	7	30%	10	9	2B
M-m-4	43	16	7	40%	8	6	
Mean	**39.5**	**16.9**	**6.3**	**43%**	**9.1**	**8.1**	
SD	**3.3**	**2.5**	**1.0**	**13%**	**0.8**	**1.1**	
Older							
O-f-1	55	9	8	20%	10	8	
O-f-2	57	9	8	50%	9	6	

O-f-3	56	12	7	10%	10	9
O-f-4	62	9	8	10%	9	5
O-m-1	60	14	7	30%	10	8
O-m-2	64	14	8	40%	10	8
O-m-3	59	9	8	10%	10	8
O-m-4	56	16	8	30%	10	8
Mean	**58.6**	**11.5**	**7.8**	**25%**	**9.8**	**7.5**
SD	**3.2**	**2.9**	**0.5**	**15%**	**0.5**	**1.3**

*2B is the required and an average grade of National Test of Oral Putonghua Proficiency for teachers or students at teachers colleges. 2A is the above-average grade.

Appendix B. Percent use of Mandarin in different situations by the three age groups of QZ Southern Min-Mandarin bilinguals

	With grand-parents	With parents	With children	With relatives	At home	At school	With co-workers	With friends	Other occasions
Younger									
Y-f-1	0	20	-	40	30	90	-	70	80
Y-f-2	0	30	-	30	30	90	90	90	60
Y-f-3	0	10	-	20	20	90	90	90	90
Y-f-4	0	50	-	20	50	90	-	60	70
Y-m-1	0	20	-	20	20	90	-	40	90
Y-m-2	0	10	-	40	20	60	70	60	70
Y-m-3	0	10	-	40	10	60	70	40	50
Y-m-4	0	10	-	30	10	80	-	50	90
Mean	**0**	**21.3**	-	**30.0**	**23.8**	**81.3**	**80.0**	**62.5**	**75.0**
SD	**0**	**12.7**	-	**8.7**	**12.2**	**12.7**	**10**	**18.5**	**14.1**
Mid-age									
M-f-1	0	0	80	10	50	80	80	50	80
M-f-2	0	10	-	50	60	90	90	80	80

M-f-3	0	10	80	30	40	80	90	80	80
M-f-4	0	0	80	20	40	50	20	20	40
M-m-1	0	30	-	50	50	50	30	30	40
M-m-2	0	0	70	10	10	40	10	10	40
M-m-3	0	0	-	0	10	50	80	10	60
M-m-4	0	0	30	0	20	50	50	20	50
Mean	**0**	**6.3**	**68.0**	**21.3**	**35.0**	**61.3**	**56.3**	**37.5**	**58.8**
SD	**0**	**10.6**	**21.7**	**20.3**	**19.3**	**18.9**	**32.9**	**29.2**	**18.9**
Older									
O-f-1	0	0	40	0	20	70	20	20	50
O-f-2	0	0	50	50	40	60	20	30	60
O-f-3	0	0	10	10	10	20	0	0	10
O-f-4	0	0	0	0	0	60	40	20	30
O-m-1	0	0	30	20	20	50	40	40	50
O-m-2	0	0	30	30	30	40	50	30	40
O-m-3	0	0	10	0	10	60	10	20	40
O-m-4	0	0	30	0	20	80	50	20	60
Mean	**0**	**0**	**25.0**	**13.8**	**18.8**	**55.0**	**28.8**	**22.5**	**42.5**
SD	**0**	**0**	**16.9**	**18.5**	**12.5**	**18.5**	**18.9**	**11.6**	**16.7**