Prosodic focus with and without post-focus compression: A typological divide within the same language family?

YI XU, SZU-WEI CHEN, and BEI WANG

Abstract

There is increasing evidence that many languages encode prosodic focus not only with phonetic variations in the focused component itself, but also with a reduction of pitch range and intensity of the post-focus components, a strategy known as post-focus compression (PFC). However, evidence is also emerging that in many other languages prosodic encoding of focus is markedly different, suggesting that PFC might be related to factors such as the presence of lexical tone, stress or the availability of morphosyntactic means of signaling focus. The current study investigated the production and perception of focus in Taiwanese, Taiwan Mandarin and Beijing Mandarin, three languages/dialects that are all tonal and that have similar morphosyntactic means for indicating focus. Results showed clear evidence of PFC in Beijing Mandarin but lack of it in Taiwanese and Taiwan Mandarin, suggesting that PFC is independent of the factors mentioned above. Most interestingly, Taiwan Mandarin seems to have lost PFC due to close contact with Taiwanese despite its effectiveness as demonstrated by the perceptual experiment. The new findings, taken together with other recent finding about prosodic focus, seem to suggest that PFC is a "hard-to-evolve" prosodic feature that may have a single historical origin. Thus there is a need for large scale experimental research to explore the cross-linguistic distribution of PFC, so as to broaden our understanding of not only prosodic typology, but also language contact, bilingualism and language evolution in general.

1. Introduction

Focus is a communicative function that serves to highlight or emphasize a particular part of an utterance. Focus can be realized by morphosyntactic

The Linguistic Review 29 (2012), 131–147
DOI 10.1515/tlr-2012-0006
0167–6318/12/029-0131

means, e.g., clefting, as in "it is XX that . . .". But emphasis can also be achieved by means of prosody, and much research has taken place to investigate the prosodic realization of focus in various languages. An important issue about prosodic focus is whether and how it can be realized in a tone language. Lexical tones, which use F_0 as their major acoustic carrier, are potentially in conflict with focus, which also uses F_0 variations as its major acoustic correlate. Thus, it is possible that the presence of lexical tones in a language would prevent F_0 from being used to encode focus (Kügler and Skopeteas 2007). However, research on Mandarin, a tone language, and Japanese, a pitch accent language, has shown that focus can be realized by F_0 variations that are independent of those due to lexical contrasts (Xu 1999; Ishihara 2002).

A general finding about prosodic focus is that a focused component is given expanded pitch range, lengthened duration, increased intensity and boosted upper spectral energy (Cooper et al. 1985; de Jong 2004; Heldner 2003). There is increasing evidence, however, that focus is also realized by changes in the post-focus components, particularly in terms of reduced pitch range (Ishihara 2002; Liu and Xu 2005; Pierrehumbert and Beckman 1988; Xu 1999; Xu and Xu 2005). Furthermore, it is found that in English yes-no questions, post-focus pitch range is compressed but raised (Eady and Cooper 1986; Liu and Xu 2007), suggesting that compression rather than lowering is the most characteristic change to the post-focus pitch range. We may thus refer to this type of focus realization as post-focus compression (PFC). There is also evidence that PFC is a highly effective perceptual cue (Botinis et al. 1999; Liu and Xu 2005; Rump and Collier 1996; Xu et al. 2004). Beside English and Mandarin, evidence of PFC is seen in many other languages, though not always in explicit forms, including German (Féry and Kügler 2008), Greek (Botinis et al. 1999), Dutch (Rump and Collier 1996), Swedish (Bruce 1982), Japanese (Ishihara 2002) and Korean (Lee and Xu 2010).

The finding of PFC in Mandarin might suggest that it is a common feature of the Chinese language family. However, Pan (2007) has reported that in Taiwanese, which is a branch of Southern Min Chinese (Min Nan Hua) spoken in Taiwan, duration is more consistently changed than F_0 in focus. However, duration increase under focus also consistently occurs in Beijing Mandarin (Chen and Gussenhoven 2008; Liu and Xu 2005; Xu 1999), English (Cooper et al. 1985; Xu and Xu 2005), Dutch (Sluijter and van Heuven 1995) and Swedish (Heldner and Strangert 2001). In addition, since no perception tests have been conducted, it is not clear how effectively focus is encoded in Taiwanese.

Before probing into the prosodic encoding of focus in these languages/dialects, a brief overview of the historical backgrounds and linguistic environments of these three languages/dialects are provided. Beijing Mandarin is the local dialect of Beijing, and its phonological system is also the basis of Stan-

dard Chinese, the official language of mainland China. Taiwan Mandarin is a variant of Beijing Mandarin spoken in Taiwan. Historically, Mandarin in Taiwan was spoken only by those who came from Mainland China. Although once homogeneous with Standard Chinese, at least by definition, it now presents noticeable differences in vocabulary, grammar (Cheng 1985) and pronunciation (Fon et al. 2004) from its mainland counterpart due to its special historical background and the multi-cultural environment in which it has developed.

There are four main ethnic groups in Taiwan: aborigines (1.7%), Hoklo (73.3%), Hakka (12%), and Mainlanders (13%) (Huang 1993: 21). Language contact has been taking place constantly in the everyday lives of these populations, which has inevitably led to societal bilingualism.[1] During the Japanese colonial period (1895–1945), most people in Taiwan were bilinguals speaking both Japanese and their own native languages. After the Second World War, Mandarin was strongly promoted by the Nationalist government until the 1980s, and it took over Japanese as the lingua franca and soon became the official and dominant language in Taiwan. Taiwanese is a branch of Southern Min Chinese, Hokkien (Min Nan), which is closely related to Amoy dialect (Lewis 2009). Over 70% of the people in Taiwan speak Taiwanese, though it was forbidden in public places and the media under the martial law until 1987. As a result, most people in Taiwan are now bilinguals, fluent in both their native languages and Taiwan Mandarin. Since Taiwanese is the most widely spoken native language, in most situations, Mandarin is used in formal occasions such as at school or in broadcasting, while Taiwanese is more dominant in daily conversations. Over the years, Taiwan Mandarin has acquired many Taiwanese features in both syntax (Cheng 1985) and phonology (Zheng 1999). Nevertheless, given the functional importance of focus, we expected little change in focus realization in Taiwan Mandarin from that in Beijing Mandarin, so it could serve as a control for highlighting the difference in focus realization between Mandarin and Taiwanese.

The goal of the present study is to make a systematic cross-linguistic comparison of prosodic focus in Taiwanese, Taiwan Mandarin and Beijing Mandarin by examining the production and perception of focus by monolingual Beijing Mandarin, monolingual Taiwan Mandarin, and monolingual Taiwanese speakers, as well as bilingual Mandarin-Taiwanese speakers. The study was designed to address three issues: (1) How does focus affect F_0, intensity, and duration in the three language/dialects? (2) Do Taiwanese speakers manipulate duration more consistently than F_0 in expressing focus? (3) Perceptually, what is the most crucial element in conveying focus?

1. Societal bilingualism occurs when in a given society two or more languages are spoken (Appel and Muysken 2005).

2. Production experiment

2.1. Method

2.1.1. Materials. The target sentence is made up of three words consisting of syllables with identical underlying tones (tone 1, high level) in both Mandarin and Taiwanese, as shown in Table 1. Although the lexical items in this sentence are the same, for Taiwanese, the tones of the 1st, 3rd, and 4th syllables change into tone 7 (mid) due to a tone sandhi rule (Chen 2000).

To elicit focus on different words in the sentence, a picture illustrating "Mother is stroking the kitten" was prepared. And a set of precursor questions, each asking about a specific aspect of the picture, were used to elicit one of the four types of focus: none (neutral focus), initial focus (on word 1), medial focus (on word 2), and final focus (on word 3) as shown in Table 2. The target sentences and their precursor questions were randomized and repeated five times. Thus, there were 4 foci × 5 repetitions = 20 sentences for each language.

2.1.2. Subjects. Four groups of 8 speakers, each with 4 males and 4 females, participated as subjects. Each monolingual speaker recorded one set of data,

Table 1. Target sentences in Taiwanese and Mandarin. The numbers in the transcription indicate the underlying tone.

	WORD 1	WORD 2	WORD 3
Characters & gloss	媽媽 'mother'	摸 'stroke'	貓咪 'kitty'
Taiwanese transcription	/ma1ma1/	/boŋ1/	/niau1mi1/
Mandarin transcription	/ma1ma1/[2]	/mo1/	/mao1mi1/

Table 2. Precursor questions in Mandarin for eliciting four types of focus.

Focus	Precursor questions	Pinyin in Mandarin	English translation
None	圖中你看見什麼?	tu2 zhong1 ni3 kan4 jian4 shen2 mo	What do you see in the picture?
Initial	誰在摸貓咪?	shei2 zai4 mo1 mao1 mi1	Who is stroking the kitty?
Medial	媽媽對貓咪做什麼?	ma1 ma1 dui4 mao1 mi1 zuo4 shen2 mo	What is Mom doing to the kitty?
Final	媽媽在摸什麼?	ma1 ma1 zai4 mo1 shen2 mo	What is Mom stroking?

2. The word for mother in Mandarin, "mama" is typically said with the neutral tone on the second syllable. For the sake of this study, we used a version of the word with Tone 1 (High-level) on both syllables. Our subjects had no problem saying this version of the word.

either in Mandarin or in Taiwanese, while bilingual speakers recorded two sets of data, in both Mandarin and Taiwanese.

Group 1: 8 monolingual Beijing Mandarin speakers, aged 18 to 30.
Group 2: 8 monolingual Taiwan Mandarin speakers, aged 26 to 60.
Group 3: 8 monolingual Taiwanese speakers, aged 28 to 58.
Group 4: 8 bilingual speakers, aged 28 to 32.

To guarantee minimal dialectal variability in Taiwanese, only native speakers born and raised in Kaohsiung and Tainan participated as subjects. None of them reported having any speech or hearing disorders.

2.1.3. Recording procedure. Each recording session took place in a quiet room. For the recording sessions in Taiwan, the speech was directly digitized into a SONY Hi-MD (MZ-RH1) recorder, using a unidirectional microphone (Audio-Technica AT 9470) placed about 5–10 inches from the subject's lips. For the recording sessions in Beijing, the speech was digitized into a computer by a 24 Bits/96K Firewire Recording System (PreSonus Firebox), using a condenser microphone (Rode NT1-A). During each trial, the experimenter read aloud the precursor question (or played the pre-recorded question in the case of Beijing Mandarin), and the subject read aloud the target sentence as an answer to the question. Each subject went through a number of practice trials until s/he was familiar with the procedure. Each subject recorded the questions in five sub-sessions, with a 5-second break in between. Within each sub-session, the order of the precursor questions was randomized.

2.1.4. F_0 extraction. The extraction of F_0 contours was done with a procedure that combines automatic vocal pulse marking by Praat (Boersma 2001) and manual rectification using a custom-written Praat script (Xu 2005–2011). When the script was run, two windows, one with the waveform and pulse markings and the other with TextGrid together with the spectrogram, were displayed. The vocal pulse markings generated by Praat were then manually rectified in the pulse window for errors such as missed or double marked vocal cycles. Segmentation of the syllables was done manually in the TextGrid window. The script then generated a smoothed F_0 contour for each sentence, and computed mean F_0, mean intensity and duration of each syllable.

2.2. Analysis

Figure 1 displays time-normalized mean F_0 contours produced by all speaker groups in four focus conditions. Each curve is an average of 40 repetitions by 8 speakers. The mean F_0 contours of Taiwanese, by both monolingual and bilingual speakers, show very little differences across the four focus conditions

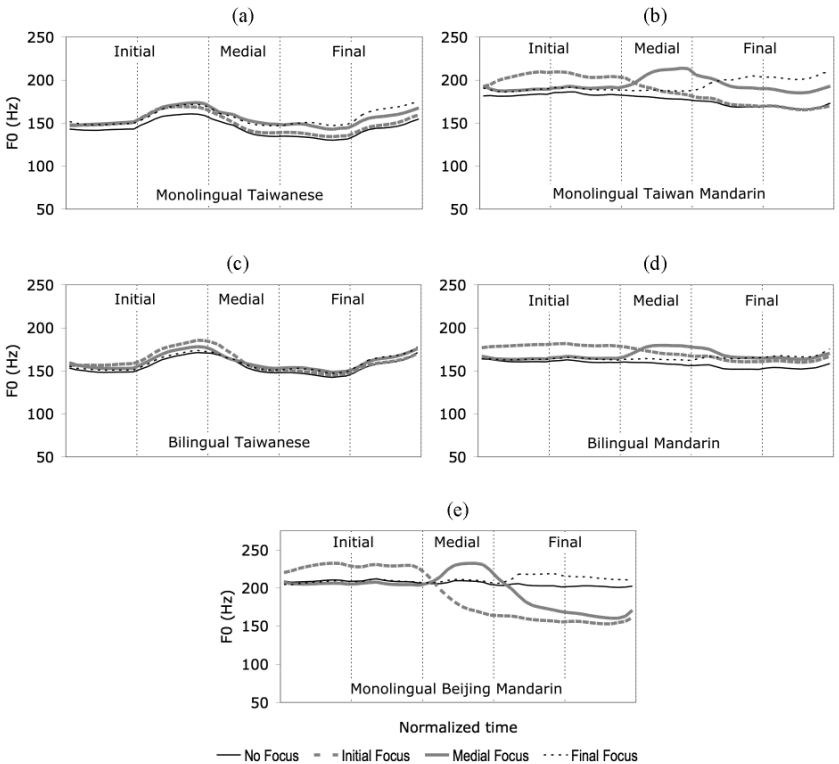

Figure 1. Time-normalized mean F_0 contours produced by 4 speaker groups. Each curve is an average of 40 repetitions by 8 speakers. The vertical lines mark the syllable boundaries.

(Figures 1a, 1c). Larger differences can be seen in the F_0 contours of Taiwan Mandarin speakers, especially those by monolingual speakers (Figures 1b, 1d). However, in none of these cases does post-focus F_0 in initial and medial focus sentences go below the F_0 of the corresponding words in the no focus condition. In contrast, post-focus F_0 is substantially lowered in the case of Beijing Mandarin (Figure 1e).

Figure 2 displays the differences in mean F_0, mean intensity and duration between the on-focus words and their no-focus counterparts. Each bar shows a value resulting from subtracting the no-focus mean from the on-focus or post-focus mean. It can be seen that on-focus raising of F_0 (Figure 2a), intensity (Figure 2c) and duration (Figure 2e) is produced by all speaker groups, and often more by speakers from Taiwan than by those from Beijing. Two-way (speaker group, focus location: initial, medial) mixed ANOVAs showed a significant effect of speaker group on on-focus duration change ($F[4,35] = 4.09$,

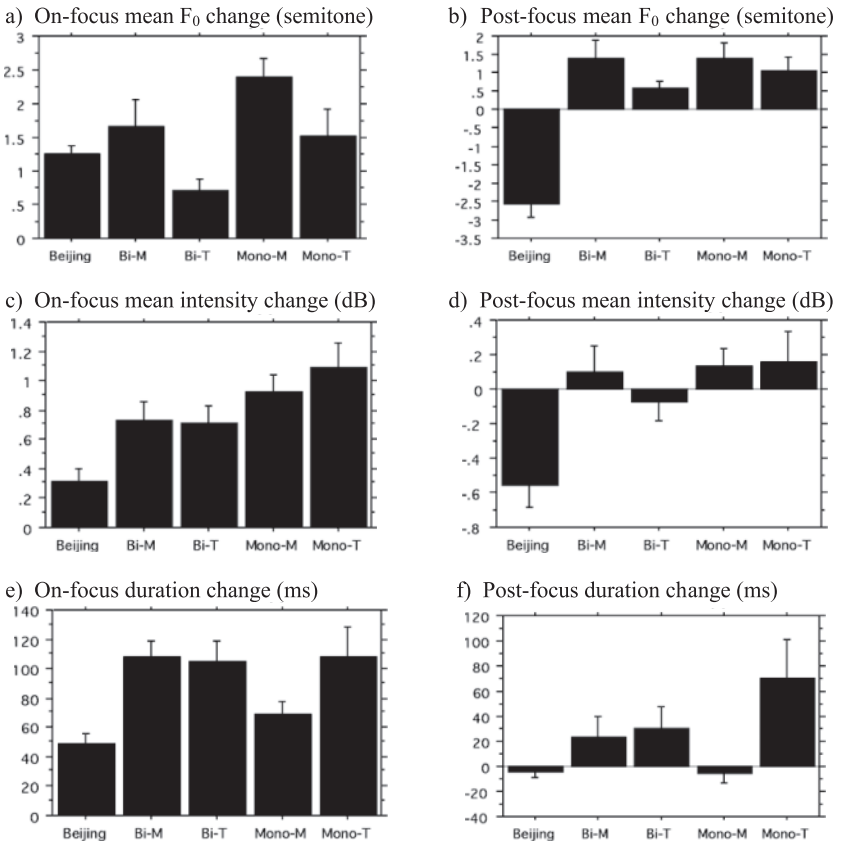

Figure 2. Differences in mean F_0, intensity and duration between on- and post-focus sentences and the no-focus ones (with standard error bars) with data from Beijing Mandarin (Beijing), bilingual Taiwan Mandarin (Bi-M), bilingual Taiwanese (Bi-T), monolingual Taiwan Mandarin (Mono-M) and monolingual Taiwanese (Mono-T) speakers.

$p < .01$), but not on on-focus change in mean F_0 or mean intensity. A Student-Newman-Keuls post-hoc analysis showed that monolingual and bilingual Taiwanese and bilingual Taiwan Mandarin speakers all had significantly longer on-focus duration than Beijing Mandarin speakers.

In contrast, only Beijing Mandarin speakers produced post-focus lowering of F_0 and intensity (Figures 2b and 2d). Another set of two-way mixed ANOVAs showed significant effect of speaker group on post-focus change in mean F_0 ($F[4,35] = 12.32$, $p < .0001$), but not on mean intensity ($F[4,35] = 2.515$, $p = 0.059$). There was a significant interaction between speaker group and focus location ($F[4,35] = 5.103$, $p < .001$). This is due to the fact that the difference

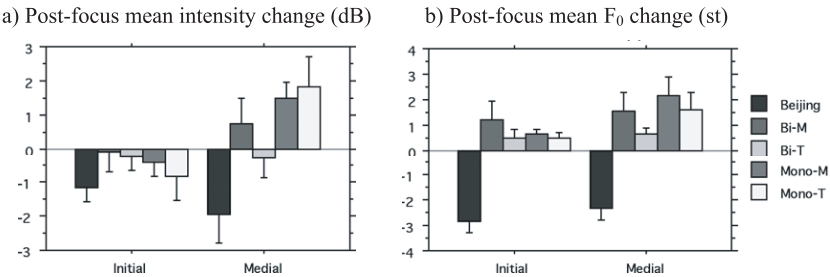

Figure 3. Interaction of speaker group and focus location on post-focus change in mean intensity (a) and mean F_0 (b).

between Beijing Mandarin and other speaker groups is clear only for medial focus, as shown in Figure 3a. For mean F_0, there was no interaction between speaker group and focus location. As shown in Figure 3b, Beijing Mandarin had lower mean F_0 than the other speaker groups for both initial and medial focus. There was an effect of speaker group on post-focus change in duration ($F[4,35] = 2.81$, $p < .05$). But as can be seen in Figure 2f, post-focus duration seems to be lengthened rather than shortened by speakers from Taiwan except monolingual Taiwan Mandarin speakers, as supported by a significant interaction between speaker group and focus location on duration ($F[4,35] = 3.275$, $p < .05$). A Student-Newman-Keuls post-hoc analysis showed significant differences in post-focus duration between Beijing Mandarin and monolingual Taiwanese speakers and between monolingual Taiwan Mandarin and monolingual Taiwanese speakers.

2.3. Discussion

Acoustically, the main difference among the three language/dialect groups is in terms of presence and absence of post-focus compression of F_0 and, to a lesser extent, intensity: In Beijing Mandarin, F_0 and intensity of post-focus words are substantially lowered, while in Taiwanese and Taiwan Mandarin, spoken by both monolingual and bilingual speakers, post-focus compression is either totally absent (F_0) or at least for medial focus (intensity). At the same time, speakers in all groups increased F_0, intensity and duration of on-focus words. There is virtually no reduction of the duration of the post-focus words, and in fact post-focus duration is *increased* in Taiwanese by monolingual speakers (No other significant differences per post-hoc analysis). Thus they increased the duration of all syllables whenever there is a focus anywhere in the sentence, which does not seem to be an effective way of encoding focus, as will be seen in the perception results to be presented next.

3. Perception experiment

3.1. Method

3.1.1. Materials. All the stimuli came from the sentences recorded in the production experiment. For each language group, three speakers were selected based on their mean standard deviation of all F_0 points across the four focus conditions: those with a) maximum, b) minimum and c) median standard deviations. All 5 tokens recorded for each of these speakers were used. So, for each language group, there were 4 foci × 5 repetitions × 3 speakers = 60 tokens.

3.1.2. Subjects. Four groups of listeners, as shown below, participated as subjects, each listening to focus samples from their own matched language groups. The bilingual group listened to both Taiwanese and Mandarin stimuli produced by bilingual speakers. Listeners had no self-reported speech or hearing disorders. Their age range was comparable to the age range of those in the production experiment.

Group 1: 11 monolingual Beijing Mandarin speakers, 6 females, 5 males, aged 18–23.

Group 2: 10 monolingual Taiwan Mandarin speakers, aged 25 to 40.

Group 3: 10 monolingual Taiwanese speakers, aged 46 to 60.

Group 4: 10 bilingual speakers, aged 28 to 52.

3.1.3. Listening procedure. The perception experiment was run with ExperimentMFC in Praat. Subjects were asked to listen to the sentence, "Mama bong niaumi" (Taiwanese) or "Mama mo maomi" (Mandarin) and judge which of the three words, or none of the words, were emphasized. They had five practice trials before the real trials without feedback on the correctness of their answers so as not to introduce any bias. In each trial, the stimulus sentence was played once.

3.2. Results and analyses

Figure 4 illustrates the average accuracy rate of focus perception in each language group. It can be seen that the overall focus recognition rate is higher for Beijing listeners than for Taiwan listeners. A two-way (speaker group, focus type: neutral, initial, medial, final) mixed ANOVA showed a significant effect of speaker group ($F[4, 46] = 14.73$, $p < .0001$), but no effect of focus.

Table 3 shows the confusion matrix of focus perception. There was a significant interaction of speaker group and focus type ($F[12, 138] = 2.11$, $p < 0.05$). This is due to the fact that the large differences between Beijing and

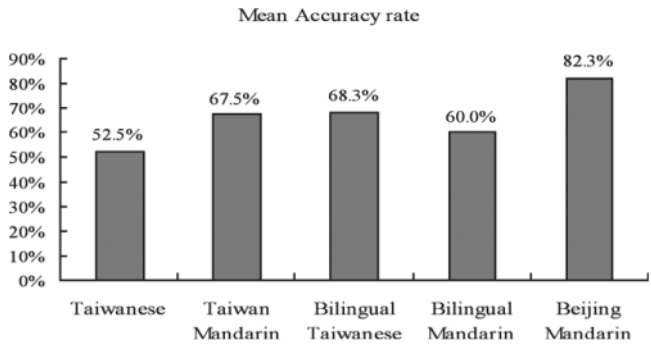

Figure 4. Focus perception by four language groups in terms of mean accuracy rate.

Taiwan listeners occur only for initial and medial focus, where compression of post-focus F_0 and intensity is possible. For final and neutral focus, Beijing listeners did not do much better than the other listeners.

3.3. Discussion

The results of the perception tests demonstrate the effectiveness of post-focus compression for encoding focus. This is shown by the fact that without such compression in Taiwanese and Taiwan Mandarin, focus recognition rate is substantially lower (50.7–73.3% correct identification, which is nevertheless well above the chance level of 25% with 4 alternative focus judgments, thanks probably to the phonetic enhancement of the on-focus words, c.f. Wu and Xu 2001 for Cantonese) than in Beijing Mandarin (66.7–90.9%), and by the fact that even for Beijing Mandarin, when such compression is not possible in final focus, the recognition rate (66.7%) is similar to Taiwanese and Taiwan Mandarin.

4. General discussion

The results of the present study largely answered our research questions raised in the introduction. The acoustic analyses and perceptual tests demonstrate that there are clear differences in the manner of prosodically realizing focus between Taiwanese and Taiwan Mandarin on the one hand and Beijing Mandarin on the other. All speakers raised on-focus F_0, intensity and duration, but only Beijing Mandarin speakers consistently produced post-focus lowering of F_0

Table 3. Confusion matrix of focus perception (%) by 4 groups of listeners listening to utterances by their own language group. Bold face indicates correct focus identification.

Listener group	Original focus	Heard as				Listener group	Original focus	Heard as			
		None	Initial	Medial	Final			None	Initial	Medial	Final
Taiwanese	None	**50.7**	22.0	16.0	10.7	Bilingual Taiwanese	None	**71.3**	14.7	10.7	3.3
	Initial	29.3	**59.3**	5.3	6.7		Initial	27.3	**61.3**	10.0	1.3
	Medial	22.0	19.3	**45.3**	14.0		Medial	18.0	2.7	**73.3**	6.0
	Final	30.7	6.7	8.7	**54.7**		Final	23.3	0.7	8.0	**68.0**
Taiwan Mandarin	None	**66.0**	13.3	18.0	2.7	Bilingual Mandarin	None	**72.0**	11.3	6.7	10.0
	Initial	32.0	**63.3**	4.0	0.7		Initial	35.3	**52.0**	9.3	3.3
	Medial	18.7	4.7	**73.3**	3.3		Medial	25.3	6.0	**63.3**	5.3
	Final	21.3	2.7	8.0	**68.0**		Final	32.0	4.7	10.7	**52.7**
Beijing Mandarin	None	**78.8**	6.7	7.3	7.3						
	Initial	7.9	**90.9**	1.2	0						
	Medial	6.7	0.6	**92.7**	0						
	Final	27.9	0	5.5	**66.7**						

and intensity. Monolingual and bilingual Taiwanese and bilingual Taiwan Mandarin did have greater on-focus duration increase than Beijing Mandarin, but they also increased post-focus duration.

The perception results suggest that the most effective device for signaling focus is reduction of post-focus F_0 and intensity. This is consistent with previous findings about focus perception (Botinis et al. 1999; Liu and Xu 2005; Rump and Collier 1996). Although focus can be perceived by speakers of Taiwanese and Taiwan Mandarin, the recognition rate is significantly lower than that of Beijing Mandarin. This indicates that the focus-related duration changes are not as effective as pitch changes.

Perhaps the most unexpected finding of the present study is that Taiwan Mandarin, which is closely related to Beijing Mandarin, realizes focus in a manner very similar to Taiwanese. This finding may have significant implications. First, the fact that PFC can be either present or absent in such closely related dialects as Beijing and Taiwan Mandarin may suggest that its adoption is largely independent of the tonal typology of a language, or whether there are morphosyntactic means of signaling focus in the language.[3] Secondly, given the origin and historical path of Taiwan Mandarin, the absence of PFC is likely a "loss" as a result of being in close contact with Taiwanese, mainly through bilingualism, which has been highly common in Taiwan since 1949 (Tsao 2000).

If indeed Taiwan Mandarin has lost PFC through close contact with Taiwanese, a natural question would arise: why is it that Beijing Mandarin has PFC in the first place? There seem to be at least three possibilities: (1) it emerged locally in the language, (2) it entered there through language contact, just as it got lost in Taiwan Mandarin through language contact, and (3) it was inherited from a proto-language.

Regarding the first possibility, it was found recently that PFC is also absent in many other languages that vary in tonality (tonal or non-tonal), word prosody (presence or absence of lexical stress), geographical or genetic affinity (from close to unrelated) to Mandarin, including Cantonese (Wu and Xu 2010), Deang, Wa and Yi (Nuosu) (Wang, et al., 2011), and Hausa, Chichewa, Wolof,

3. One may argue that, because Cantonese, Taiwanese and Taiwan Mandarin (and perhaps other Southern Chinese dialects) all lack lexical stress, the lack of PFC in these languages is due to the lack of lexical stress. It could also be argued, as suggested by one reviewer, that although known PFC languages like Japanese and Korean also do not have lexical stress, their pitch accent systems are in many ways similar to a stress system, e.g., culminativity, abstract relation of prominence, etc., lexical stress still cannot be ruled out as a possible trigger for PFC. We would say that our current state of the knowledge cannot completely rule out any of these possibilities. And that is why there is a need to investigate prosodic focus in many more languages.

Northern Sotho and Buli, etc. (reviewed by Zerbian et al., 2010). The fact that PFC did not arise automatically in these languages at least suggests that PFC may not easily emerge in a language.

Regarding the second possibility, historically, Northern China was in close contact with many non-Chinese speaking populations, in particular, Mongolian and Manchurian, who ruled China during the Yuan (1271–1368 AD) and Qing (1644–1912 AD) dynasties. As a result of such contact, there has been much influence of those languages on Mandarin (Chappell 2001; LaPolla 2001; Wadley 1996). Both Mongolian and Manchurian are Altaic, a hypothetic language family that includes the Turkic, Mongolic and Tungusic languages (Georg et al. 1999), and Korean and Japanese according to some scholars (Georg et al. 1999). Interestingly, there is evidence that many Altaic languages have PFC (Japanese: Ishihara 2002; Korean: Lee and Xu 2010; Turkish: Ipek 2011; Uygur: Wang et al., 2011). Thus it is possible that PFC was spread into Mandarin from Altaic languages through language contact.

There are two difficulties with the spreading account, however. The first is that there is evidence that many European languages also have PFC (Botinis et al. 1999; Bruce 1982; Féry et al. 2008; Rump and Collier; Xu and Xu 2005), but it is unlikely that they all once had close contact with Altaic languages.[4] The second difficulty is that so far there is evidence only for the loss of PFC through language contact, e.g., English without prosodic focus when spoken as a second language (Swerts and Zerbian 2010), or Mandarin without PFC after contact with a non-PFC language (present data, and Wang, 2011), but no report of a language gaining PFC through language contact, e.g., spreading PFC from Mandarin to Taiwanese (present study), from English to Cantonese (Wu and Chung, 2011), or from Mandarin to Deang (a non-tonal language in the Mon-Khmer language family. See: Wang et al., 2011).

The third possibility, namely, PFC is inherited from a proto-language, is at first glance the least likely, because it would entail that Mandarin is a descendant of Altaic languages and that there is a common ancestor to both European and Altaic languages. The first entailment would exceed even the strongest claims about the Altaic influence on Chinese (Chappell 2001; LaPolla 2001; Wadley 1996), and the second entailment is reminiscent of controversial proposals such as the Eurasiatic (Greenberg 2000) or the Nostratic (Bomhard 2008) macrofamilies (see Xu 2011 for further speculations along this line). On the other hand, if the unspreadability of PFC is supported by further research, the inheritance hypothesis should be taken seriously.

4. It is possible, as suggested by one reviewer, that PFC in European languages is indigenous or has another source. Of course, if this is the case, there is no need to consider the spreading hypothesis.

5. Conclusion

The present findings show that PFC, i.e., post-focus compression of pitch range and intensity, as a prosodic device for signaling focus can be either present or absent even in very closely related languages. Such a disparity thus seems to be independent of whether the language is tonal, such as Taiwanese, Taiwan Mandarin and Beijing Mandarin, or whether there are morphosyntactic means to indicate focus, as they are present in both Mandarin and Taiwanese. The case of Taiwan Mandarin is especially intriguing, as it is phonetically very similar to Beijing Mandarin, and yet its focus realization is more similar to Taiwanese, with which it has been in close contact for several generations. The present data have also once more demonstrated the perceptual benefit of PFC. Its presence in initial and medial focus in Beijing Mandarin lead to over 90% focus recognition, whereas the lack of it in final focus in Beijing Mandarin and in all types of focus in Taiwanese and Taiwan Mandarin lead to less than 75% of focus recognition. Although 75% or even lower is still well above chance, the improvement of focus perception with PFC is probably sizable enough for PFC to be maintained in many languages once it is in place.

These findings, when considered in conjunction with other recent findings, suggest the possibility that PFC has a single historical origin rather than having developed separately due to the tonal, accentual, and morphosyntactic characteristics of individual languages. Such a hypothesis is of course highly speculative at this moment. To test it, a much larger-scale typological investigation of the world's languages is needed. If sufficient support is found for the hypothesis, the implications could be profound for many areas, including linguistic typology, linguistic as well as biological human evolution and interface of prosody with other aspects of speech. Finally, our data also suggest the importance of using systematic experimental controls in these investigations, such as eliciting focus with context, using identical target sentences, taking measurements from both on-focus and off-focus syllables, and the inclusion of a no focus condition as the base line.

Acknowledgement

The experiments on Beijing Mandarin were partly supported by "Project 211" of Minzu University of China financed by Ministry of Education of P. R. China to the third author.

University College London
yi.xu@ucl.ac.uk

National Chung Cheng University
chen.szuwei@gmail.com

Minzu University of China
bjwangbei@googlemail.com

References

Appel, Rene & Muysken, Pieter. 2005. *Language contact and bilingualism*. Amsterdam: Amsterdam University Press, Amsterdam Academic Archive.

Boersma, Paul. 2001. *Praat, a system for doing phonetics by computer*. Glot International 5:9/10:341–345.

Bomhard, Allan R. 2008. *Reconstructing Proto-Nostratic: Comparative Phonology, Morphology, and Vocabulary*. Leiden: Brill.

Botinis, Antonis, Marios Fourakis & Barbara Gawronska. 1999. Focus identification in English, Greek and Swedish. *XIVth International Congress of Phonetic Sciences*, vol. 2, 1557–60, San Francisco.

Bruce, Gosta. 1982. Developing the Swedish intonation model. *Lund University, Dept. of Linguistics Working Papers* 22. 51–116.

Chappell, Hilary. 2001. Language contact and areal diffusion in Sinitic languages: Problems for typology and genetic affiliation. In A. Y. Aikhenvald and R. M. W. Dixon (eds.), *A real diffusion and genetic inheritance: problems in comparative linguistics*, 167–194.

Chen, Matthew Y. 2000. *Tone sandhi: Patterns across Chinese dialects*. Cambridge, UK: Cambridge University Press.

Chen, Yiya & Carlos Gussenhoven. 2008. Emphasis and tonal implementation in Standard Chinese. *Journal of Phonetics* 36. 724–746.

Cheng, Robert L. 1985. A Comparison of Taiwanese, Taiwan Mandarin, and Peking Mandarin. *Language* 61(2). 352–377.

Cooper, William, Stephen Eady, Pamela Mueller. 1985. Acoustical aspects of contrastive stress in question-answer contexts. *Journal of the Acoustical Society of America* 77(6). 2142–2155.

Jong, Kenneth de. 2004. Stress, lexical focus, and segmental focus in English: patterns of variation in vowel duration. *Journal of Phonetics* 32. 493–516.

Eady, Stephen J., & William E. Cooper. 1986. Speech intonation and focus location in matched statements and questions. *Journal of the Acoustical Society of America* 80. 402–415.

Féry, Caroline & Frank Kügler. 2008. Pitch accent scaling on given, new and focused constituents in German. *Journal of Phonetics* 36. 680–703.

Fon, Janice, Wen-Yu Chiang, & Hintat Cheung. 2004. Production and perception of two dipping tones (T2 and T3) in Taiwan Mandarin. *Journal of Chinese Linguistics* 32 (2). 249–280.

Greenberg, Joseph Harold. 2000. Indo-European and Its Closest Relatives: The Eurasiatic Language Family. Volume 1, Grammar. Stanford: Stanford University Press.

Georg, Stefan, Peter A. Michalove, Alexis Manaster Ramer, & Paul J. Sidwell. 1999. Telling general linguists about Altaic. *Journal of Linguistics* 35 (1). 65–98.

Heldner, Mattias. 2003. On the reliability of overall intensity and spectral emphasis as acoustic correlates of focal accents in Swedish. *Journal of Phonetics* 31. 39–62.

Heldner, Mattias & Eva Strangert. 2001. Temporal effects of focus in Swedish. *Journal of Phonetics* 29. 329–361.

Huang, Shuanfan. 1993. *Language, Society, and Ethnic Identity*. Taipei: Crane.

Ipek, Canan. 2011. Phonetic realization of focus with no on-focus pitch range expansion in Turkish. *The 17th International Congress of Phonetic Sciences*. 140–143. Hong Kong.

Ishihara, Shinichiro. 2002. Syntax-phonology interface of wh-constructions in Japanese. *Tokyo Conference on Psycholinguistics*. 165–189. Tokyo.

Kügler, Frank & Stavros Skopeteas. 2007. On the universality of prosodic reflexes of contrast: The case of Yucatec Maya. *International Congress of Phonetic Sciences (ICPhS)*, Saarbrücken. 16. 1025–1028.

LaPolla, Randy J. 2001. The Role of Migration and Language Contact in the Development of the Sino-Tibetan Language Family. In R. M. W. Dixon and A. Y. Aikhenvald (eds.), *Areal Diffusion and Genetic Inheritance: Case Studies in Language Change*, 225–254. Oxford: Oxford University Press.

Lee, Yong-Cheol & Yi Xu. 2010. Phonetic Realization of Contrastive Focus in Korean. *Speech Prosody 2010*. Chicago.

Lewis, M. Paul. 2009. *Ethnologue: Languages of the World*. (Sixteenth edn.). Dallas, Tex: SIL International. Online version: http://www.ethnologue.com/.

Liu, Fang & Yi Xu. 2005. Parallel Encoding of Focus and Interrogative Meaning in Mandarin Intonation. *Phonetica* 62. 70–87.

Liu, Fang & Xu, Y. 2007. Question intonation as affected by word stress and focus in English. *The 16th International Congress of Phonetic Sciences*, 1189–1192. Saarbrücken.

Pan, Hoh. 2007. Focus and Taiwanese unchecked tones, In Lee, C., Gordon & M., Büring, D. (eds.), *Topic and Focus: Cross-linguistic Perspectives on Meaning and Intonation.* The Netherlands: Springer.

Pierrehumbert, Janet & Mary Beckman. 1988. Japanese Tone Structure, Cambridge, MA: The MIT Press.

Rump, H. H. & Collier, R. 1996. Focus conditions and the prominence of pitch-accented syllables. *Language and Speech* 39. 1–17.

Sluijter, Agaath & Vincent J. van Heuven. 1995. Effects of focus distribution, pitch accent and lexical stress on the temporal organization of syllables in Dutch. *Phonetica* 52. 71–89.

Swerts, Marc & Sabine Zerbian. 2010. Prosodic transfer in Black South African English. *Speech Prosody 2010*. Chicago.

Tsao, Feng-Fu. 2000. Taiwanized Japanese and Taiwan Mandarin-Two Case Studies of Language Contact during the Past Hundred Years in Taiwan. *Han Xue Yan Jiu [Chinese Studies]* 18. 273–97.

Wadley, Stephen A. 1996. Altaic Influences on Beijing Dialect: The Manchu Case. *Journal of the American Oriental Society* 116 (1). 99–104.

Wang, Bei, Ling Wang, Tursun Kadir. 2011. Prosodic encoding of focus in six languages in China. *The 17th International Congress of Phonetic Sciences*. 144–147. Hong Kong.

Wang, Ling. 2011. Prosodic encoding of focus in statement and question sentences: Comparing Deang, Wa, Tibetan and Mandarin. China: Minzu University of China master thesis.

Wu, Wing Li & Lisa Chung. 2011. Post-focus compression in English-Cantonese bilingual speakers. *The 17th International Congress of Phonetic Sciences*. 148–151. Hong Kong.

Wing Li Wu & Yi Xu. 2010. Prosodic Focus in Hong Kong Cantonese without Post-focus Compression. *Speech Prosody 2010.* Chicago.

Xu, Yi. 2005–2011. ProsodyPro.praat. On-line: ⟨http://www.phon.ucl.ac.uk/home/yi/tools.html⟩.

Xu, Yi. 1999. Effect of Tone and Focus on the Formation and Alignment of f0 contours. *Journal of Phonetics* 27. 55–107.

Xu, Yi, & Ching X. Xu. 2005. Phonetic realization of focus in English declarative intonation. *Journal of Phonetics* 33. 159–197.

Xu, Yi. 2007. Speech as articulatory encoding of communicative functions. *The 16th International Congress of Phonetic Sciences*, 25–30. Saarbrucken.

Xu, Y. 2011. Post-focus compression: Cross-linguistic distribution and historical origin. In *Proceedings of The 17th International Congress of Phonetic Sciences*, Hong Kong: 152–155.

Xu, Yi, Ching X. Xu & Xuejing Sun. 2004. On the Temporal Domain of Focus. *Speech Prosody 2002*, Nara, Japan.

Zerbian, Sabine. 2007. Investigating prosodic focus marking in Northern Sotho. *Focus strategies: evidence from African languages*, ed. by K. Hartmann, E. Aboh and M. Zimmermann, 55–79. Berlin, Mouton de Gruyter.

Zerbian, Sabine, Genzel, Susanne, & Frank Kügler. 2010. Experimental work on prosodically-marked information structure in selected African languages (Afroasiatic and Niger-Congo). *Speech Prosody 2010*, Chicago.

Zheng, Chin-Chin. 1999. A Contrastive Analysis on the Pronunciation between Mainland China and Taiwan: A Case Study of News Reporters. Taipei: NSC-88-2411-H-003-017.