

## Prosodic Focus in Malay without Post-focus Compression

Muhammad Shafiq Azid<sup>1\*</sup> and Yi Xu<sup>2</sup>

<sup>1</sup>*Department of Languages and Cultures, Universiti Pertahanan Nasional Malaysia, Kem Sungai Besi, 57000 Kuala Lumpur, Malaysia*

<sup>2</sup>*Department of Speech, Hearing and Phonetic Sciences, University College London, London WC1N 1PF, United Kingdom*

### ABSTRACT

Many studies across languages have recognised that focus substantially alters the prosodic structure of a sentence not only by increasing  $F_0$ , intensity, and duration of the focused words but also by compressing the range of pitch and intensity of the post-focus words. Studies, however, are still not fully clear regarding the main effects of focus on focused and post-focused words in Malay. Analyses from the present study revealed that on-focused words had significantly increased  $F_0$ , intensity, and duration, while post-focused words showed no significant lowering following the effect of focus. The outcomes of the study generalised Malay as a language without post-focus compression (PFC). These findings have implications on prosodic typology, language contact and the historical origin of Malay.

*Keywords:* Focus, Malay, post-focus compression

### ARTICLE INFO

#### *Article history:*

Received: 10 December 2019

Accepted: 30 January 2020

Published: 18 March 2020

#### *E-mail addresses:*

[mshafiq@upnm.edu.my](mailto:mshafiq@upnm.edu.my) (Muhammad Shafiq Azid)

[yi.xu@ucl.ac.uk](mailto:yi.xu@ucl.ac.uk) (Yi Xu)

\*Corresponding author

### INTRODUCTION

#### Focus

When a speaker utters a sentence, certain information in the sentence is emphasised more than the others, and this linguistic phenomenon is generally known as focus. Under a certain discourse condition, some parts of a sentence might as well get highlighted (Bolinger, 1958; Eady & Cooper, 1986; Ladd, 2008; Xu, 1999). Such a phenomenon can be manifested

morphosyntactically through clefting as in 'It is X who ...' by which the constituent X is put to focus. Additionally, prosodic alterations can also contribute to the effect of focus. Findings from various languages have shown that a focused word is characterised with a higher fundamental frequency ( $F_0$ ), a longer duration and a greater amplitude in contrast with the non-focused words (Cooper et al., 1985; Eady & Cooper, 1986; Féry & Kugler, 2008; Xu & Xu, 2005).

### Prosodic Realisation of Focus

Within the realm of autosegmental-metrical (A-M) theory of intonational phonology, the phonological constituents that accompany certain stressed syllables are known as pitch accents especially in languages where stressed syllables are prevalent like English and Dutch (Ladd, 2008). With a better understanding of stress in terms of pitch accent, changes in pitch are thought to provide major cues to the perception of stress such that pitch accent illustrates actual prominence in a sentence whereas stress describes abstract lexical properties of individual syllables.

However, dependence on this view can be problematic looking at its inability to address multiple factors and conditions that influence the variations of  $F_0$  specifically under the condition of focus. Essentially, studies have reported the effect on  $F_0$  linked to a multitude of conditions, but the A-M approach has yet to demonstrate the interaction that occurs between these discriminate factors (Xu & Xu, 2005). Hence, the A-M approach is incomprehensive at

illustrating the workings of prosody as a system in communicative terms. In lieu of the complication, it is necessary to search for a coherent approach that can compromise all these factors in understanding the effect of focus.

By employing the articulatory-functional view of speech as an alternative, various conditions affecting prosodic focus with their interacting effects can be examined through the Parallel Encoding and Target Approximation (PENTA) model. The model is a basis that assembles independent mechanisms but treats them all as an integral part of speech from which simultaneous encoding of communicative functions can be realised as a communicative system (Xu et al., 2015). The PENTA model focuses on the functions rather than the forms by viewing the role of phonology as an abstract manifestation of speech prosody. Unlike the A-M model that focuses on categorical distinctions of prosodic units, the PENTA model defines units on a functional basis.

Within this model, the prosodic focus is conceptualised as an encoding scheme that links a communicative function with articulatory parameters through several underlying prosodic components. That is prosodic focus functions to emphasise a speech unit against the others within the sentence through components like duration and intensity. Additionally, the prosodic focus is found not to be the only encoding scheme utilised to achieve focus, but the effect of post-focus compression of pitch and intensity (PFC) is also used in realising focus in several other languages (Xu et al.,

2012). At this point, viewing prosodic focus from the perspective of the PENTA model helps to explain the phenomenon in a more detailed, generalisable and predictable way. Also, the dispersed distribution of prosodic focus and PFC across languages elucidates the nature of the phenomenon as multi-faceted and language-specific.

A focused element within a sentence is largely represented by increased pitch, increased  $F_0$ , longer duration and higher amplitude in comparison to its unfocused correlates across languages. Dohen and Løvernbruck (2004) from their findings showed that in French, narrow focus (or contrastive focus) had an increased pitch as compared to broad focus. In addition, the study also showed that the duration of the focused syllables was lengthened by 33.71%, which was perceptually pertinent. Xu and Xu (2005) reported that in declarative English sentences, the narrow focus was associated with higher  $F_0$  peaks as compared to neutral-focus constituents. The increase in duration for the focused constituents has long been supported in an earlier study by Cooper et al. (1985) that showed the elongation of the emphasised words using a contrastive focus in question-answer stimuli in English. Additionally, a study done to test the reliability of overall intensity and spectral emphasis of focal accents in Swedish showed that the overall intensity of a focused constituent had a significant increase against other non-focused constituents by about 3dB (Heldner, 2003). Deductively, an on-focus element would generally be susceptible

to a higher pitch, longer duration, and greater amplitude, which lead to its marked emphasis against other constituents within a sentence. While this glaring effect of focus is rather predicted, questions aroused as to whether focus manifests in isolation or in a dynamic fashion. This entails the assumption as to whether focus has any consequential effect with respect to its neighbouring constituents or the sentence in which it is located.

By examining Chinese Mandarin, Xu (1999) suggested an asymmetric structure in illustrating the significance of focus on  $F_0$  such that there was (a) an expansion on non-final focused words; (b) a suppression or compression in the post-focus regions; and (c) a neutral shape in all other words with minimal deviation of  $F_0$  range before the focused words. Note that the pitch range adjustments structure indicates the consequential effects of focus around the on-focus region. This structure is further reinforced by a perception study on Beijing Mandarin focused words that showed 90.9% and 92.7% rate of perception in initial and medial focus conditions respectively (Xu et al., 2012). This tallies the argument made by Cooper et al. (1985) who stated that relative prominence of duration and  $F_0$  varied according to the focus location. Other studies have also shown that pre- and post-focus constituents could be influenced by the effect of focus especially in terms of their duration and  $F_0$  contour (Dohen & Løvernbruck, 2004; Jun & Lee, 1998; Xu & Xu, 2005).

From studies on prosodic elements, languages are divided into certain groups. In non-tonal languages,  $F_0$  variations have been the major approach in realising prosodic focus (Botinis et al., 1999; Rump & Collier, 1996). In contrast, tonal languages like Mandarin utilised these variations for lexical distinction and they are thought as possibly unnecessary for prosodic focus realisation.

However, an alternative view aroused as a response to such an assumption. From the observations of pitch and intensity on the post-focus elements in Beijing Mandarin, such compression was inferred as an approach to elicit focus (Jin, 1996; Liu & Xu, 2005). Nonetheless, these findings from Beijing Mandarin did not provide a forthright generalisation. The compression effect was not found in a study done on a closely-related language, Taiwanese Mandarin; rather, the features found were more similar to that of Taiwanese and Hong Kong Cantonese (Chen et al., 2009; Wu & Xu, 2010). In Hong Kong Cantonese, findings have shown that  $F_0$  of the lexical tones was not wholly manipulated for the purpose of highlighting focus because the speakers of the language are prone to keeping the feature unchanged in non-focused conditions. Interestingly, findings in light of  $F_0$  variations and focus between tonal and non-tonal languages have generated insights about the notion of PFC. While studies have long put emphasis on the phenomenon of focus, PFC as an element that emerged specifically and variably from the effect of focus should be an informative constituent in distinguishing language typological differences.

The notion of PFC has formed a pattern across languages which may provide supportive accounts on the linguistic typology, diachronic linguistics and evolutionary theory of human. Xu (2011) provided a cross-linguistic overview of PFC distribution in which he argued that, while it could be integral in underscoring focus, PFC distribution was not necessarily universal. A measure of post-focus  $F_0$  lowering, deaccenting, dephrasing and pitch range compression has suggestively drawn Dutch (Rump & Collier, 1996), French (Dohen & Løevernbruck, 2004), Finnish (Mixdorff, 2004), German (Féry & Kugler, 2008), Greek (Botinis et al., 1999), Hindi (Patil et al., 2008), Japanese (Ishihara, 2002), Korean (Lee & Xu, 2010), Egyptian, Lebanese Arabic (Chahal, 2003; Hellmuth, 2006), Tibetan, Uygur, Nanchang (Wang et al., 2011), Persian (Taheri & Xu, 2012), Turkish (Ipek, 2011) and Swedish (Bruce, 1982) into languages with PFC. On the contrary, languages without indications of PFC are found from various language families and origins such as African languages (Afroasiatic & Niger-Congo) in Buli, Chichewa, Chitumbuka, Durban Zulu, Northern Sotho, Hausa and Wolof (Rialland & Robert, 2011; Zerbian et al., 2010); Chinese languages in Yi, Deang and Wa (Wang et al., 2011); and Mayan in Yucatec Maya (Kugler & Skopeteas, 2007).

Many explanations can be made from the dispersed PFC differences which will be useful in drawing stipulations on the historical movement of languages. Indeed, three famous hypotheses have been

discussed by Xu (2011) in this regard namely (a) independent genesis or emerging in situ within a language; (b) horizontal spreading or borrowing from other languages through contact, or (c) vertical inheritance or descended from an ancient proto-language. The argument for independent genesis is seemingly faint looking at the abundant data suggesting PFC's inability to arrive in situ and its 'hard-to-evolve' nature (Xu et al., 2012) while spreading would only hold true if there had been findings showing a gain of PFC from language contact rather than just a loss of PFC from contact (Xu, 2011). At this juncture, vertical inheritance seems to be the highly endorsed hypothesis in support of the current findings because the grouping of PFC languages has shown that the feature as a means of encoding focus could possibly arrive from a common ancestor called the Nostratic superfamily. The study, in particular, presupposes this hypothesis; yet, all of these stipulations require a wealth of support for them to hold true because in the case of the hypothetical Nostratic superfamily, for instance, data from some language families are still inadequate. While it seems much safer to hold on to this hypothesis, other hypotheses should not be disregarded entirely.

### **The Case of Malay**

Many languages have been described in terms of their prosodic realisations and historical connection; however, this paper will explore Malay since not much has been described the language in the aspect of

prosodic focus especially within the PENTA model. By doing so, this paper hopes to address the aforementioned hypotheses and simultaneously provide richness for the description of the language.

In addition, there has not been any exploration of Austronesian languages on the notion of PFC given the fact that the Nostratic superfamily hypothesis argues for only Indo-European, Uralic, Altaic, Afroasiatic, Dravidian, Kartvelian, and Eskimo-Aleut languages to be in the same assemblage. Yet again, looking at the hypotheses regarding the origins of Malay speakers, which might arrive from the Southern China or Taiwan (Kern as cited in Anceaux, 1965; Bellwood, 1997; van Heine Geldern, 1966), it will be enticing to see if those assumptions would still hold true with the evidence arriving from this study. That is, if the Malay language shows strong indications of PFC, then the possibility increases for the language to have not originated from those implicated regions and the hypothesis of a proto-language origin will be vulnerable. Otherwise, the Nostratic superfamily hypothesis and the hypotheses regarding the origins of Malay speakers will still remain strongly plausible.

This paper will examine the  $F_0$  contours of contrastive focus uttered in short Malay declarative sentences which are prompted by questions. By doing so, this paper aims at answering the following queries:

1. Will there be differences in pitch, intensity, and duration between on-

- focus words with their correlates in the non-focused conditions?
2. Will there be a significant lowering pattern after the effect of focus in the post-focus words?

Based on previous studies, several predictions are assumed. First, there would be higher values on the three parameters when the words are focused on a sentence. Second, there would be no significant lowering pattern in the post-focus words under focus. Analyses from the present study will reveal the directions of these predictions.

## METHODS

### Stimuli

Four unique target sentences served as stimuli for the study. Each sentence consisted of five words to make up a simple declarative sentence to control for the effect of sentence type. The target sentences were constructed in such a way that there would be three distinct focus locations within the sentences, namely the initial-focus (first word), the medial-focus (third word) and the final-focus (fifth word). All focus locations were nouns. In addition to these three focus locations, the non-focused or neutral condition was introduced such that none of the words in the sentences was focused. This condition will serve as a baseline for comparisons.

Several studies have shown that the initial consonant of a syllable can have an influence on its  $F_0$  contour (Howie, 1974; Lehiste, 1975; Lehiste & Peterson, 1961).

To address the issue, sonorants were used in the onsets of stimuli to maintain the least obstruction and interruption on the continuous  $F_0$  contours (Xu, 1999). As per this study, sonorants were used to control for the effects across the target sentences although not in all possible syllables. In addition, to ensure the ease of segmentation during an acoustic analysis of the data, the use of glides or vowels at the start of the syllables within the target sentences were put at a minimum.

While there are many types of focus described in the literature, a type of focus that was used in this study is the contrastive focus. It is a subtype of narrow focus that can be employed to generate both syntactic and prosodic cues (Féry, 2001). Literature has indicated that contrastive focus could trigger vital prosodic marks for analysis such as  $F_0$ , duration and/or intensity via highlighting certain words or phrases within a sentence (Dohen & Løevenbruck, 2004; Féry, 2001). Since this paper examines the effect of focus on a phonological level without a particular examination of syntactic or pragmatic elements, this type of focus was selectively used for analysis.

Contrastive focus on a specific word at different focus locations was elicited using four precursor questions. Each precursor question would directly enquire a specific unit of information from the target sentence, hence a specific focus location. Targeting contrastive focus using question and answer would help in controlling for the effect of focus during the experiment. The precursor questions were randomised and repeated



across four different target sentences. An example of a target sentence with its respective precursor sentences is shown in Table 1 below.

### Participants

Twelve native speakers of Malay were recruited to participate in the experiment. Only participants who were born and raised in Malaysia with high proficiency in Malay were selected to minimise the variability between the participants. All the participants were students from universities around London with equal numbers of male and female participants. Participants' ages ranged between 23-35 years old (mean = 29 years and 5 months) and all participants reported no speech or hearing impediment. The participants were given an information and consent form prior to the actual recording sessions. The ethics of the study was approved under blanket

ethics by the Ethics Chair of University College London (project number: SHaPS-2014-YX-013).

### Procedure

The recording sessions were conducted separately for each participant in an anechoic recording room in the Division of Psychology and Language Sciences building of University College London using RODE NT1-A large-diaphragm condenser microphone with a Focusrite Scarlett 2i2 external sound interface connected to a PC. The condenser microphone was positioned about 30 cm from the participants' mouths. Before the actual recording, the participants were given printed sheets of the stimuli and asked to practice reading aloud the target sentences by highlighting foci according to the precursor questions. Then, the recording was started as the experimenter read the precursor questions to the participants.

Table 1

*The precursor questions preceding the target sentences addressing different focus locations*

Precursor Questions	Target Sentences
1. <i>Siapakah yang membeli nanas di kedai?</i> 'Who bought a pineapple at the shop?'	1. <b>Najib</b> membeli nanas di kedai 'Najib bought a pineapple at the shop'
2. <i>Apakah yang Najib beli di kedai?</i> 'What did Najib buy at the shop?'	2. Najib membeli <b>nanas</b> di kedai 'Najib bought a pineapple at the shop'
3. <i>Di manakah Najib membeli nanas?</i> 'Where did Najib buy a pineapple?'	3. Najib membeli nanas di <b>kedai</b> 'Najib bought a pineapple at the shop'
4. <i>Apakah yang sedang berlaku?</i> 'What is happening?'	4. Najib membeli nanas di kedai 'Najib bought a pineapple at the shop'

Notes: 1 = initial focus location, 2 = medial focus location, 3 = final focus location, 4 = no focus location or neutral.

The subject read aloud the target sentence as a response to the precursor questions. Whenever the experimenter noticed any sentence that was overdone, understated or inappropriately pronounced, the participants were asked to start again. Every participant read the sentences in four different blocks containing each unique sentence with four precursor questions. The participants were given one-minute break between blocks and each participant read in total 16 target sentences in one session. The experimental design can be seen in Figure 1 below. The speech signals were directly digitised onto the hard disk of a PC at a sampling rate of 22kHz and 16 bits bit width, using a software digitization program called Praat, version 5.4.12 (Boersma & Weenink, 2015).

**F<sub>0</sub> Extraction and Measurements**

The F<sub>0</sub> extraction was done using a custom-written Praat script called ProsodyPro (version 6.0.1) by Xu (2013). When the script was run, two windows were displayed with one showing the waveform and pulse markings while the other showing spectrogram and TextGrid. In the beginning, the target sentences were

segmented accordingly and labelled in the TextGrid window. The segmented data were then being manually rectified in the pulse window for any extraneous errors such as missed or irregular vocal markings. This procedure was followed by F<sub>0</sub> contour smoothing process so as to minimise random variations in the contours and subsequently maintain the accurate measurement of the location and value of F<sub>0</sub> peaks and valleys (Xu, 1999). Based on the manually labelled boundaries, several measurements were collected from the different focus locations: maximum F<sub>0</sub> (max F<sub>0</sub>), mean F<sub>0</sub>, mean intensity (henceforth, intensity) and duration as automatically generated by the ProsodyPro script. These measurements were later analysed using SPSS version 22.0 (IBM Corp., 2013).

**RESULTS AND DISCUSSIONS**

**Results**

The results obtained from the F<sub>0</sub> extraction of the Malay participants' sentence productions were transformed into a number of time-normalised mean F<sub>0</sub> contours. These contours were plotted by focus locations and different unique sentences, while the word

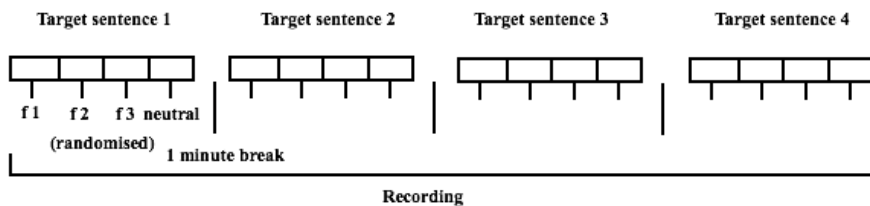


Figure 1. Experimental design of the study showing stimuli being presented in four blocks, each of which contains four focus conditions. Note. f1 = initial focus, f2 = medial focus, f3 = final focus, neutral = neutral or non-focused condition.



boundaries representing focus locations were marked with vertical dash lines. As can be seen in Figure 2, the contour shows a longer interval between initial and medial foci because the word is longer than the others in the sentence.

Figure 2 shows that there is a clear trend of increased mean  $F_0$  contours for the on-focus words in three different locations as compared to other words. The highest observable mean  $F_0$  peaks occur at the initial focus location at approximately 210 Hz. The trend in terms of mean  $F_0$  peaks for the focused words also displays a decreasing pattern towards the end of the sentence, with around 190 Hz and 180 Hz for medial and final locations respectively. Secondly, however, a general observable indication of PFC is seemingly absent, looking at the close overlap between the curves following the focus marks. A much wider gap between the

focus conditions to that of the neutral curve would alternatively suggest the possibility of PFC in the language. Nonetheless, the result of close overlap might have arrived from the effect of averaging across unique sentences. Hence, time-normalised mean  $F_0$  contours representing each unique sentence in isolation would be necessary.

Figure 3 shows that each word has a rising-falling  $F_0$  contour with one or two peaks regardless of the focus condition. General observation suggests that only Sentence 1 in Figure 3 displays PFC in the language. This is due to the consistent further lowering of  $F_0$  following the on-focus words at all focus locations as compared to their neutral correlates, such that the on-focus  $F_0$  continues to drop until it reaches a level that is lower than the neutral-focus contour (in blue). Comparing with the other sentences, Sentence 2 does not show observable

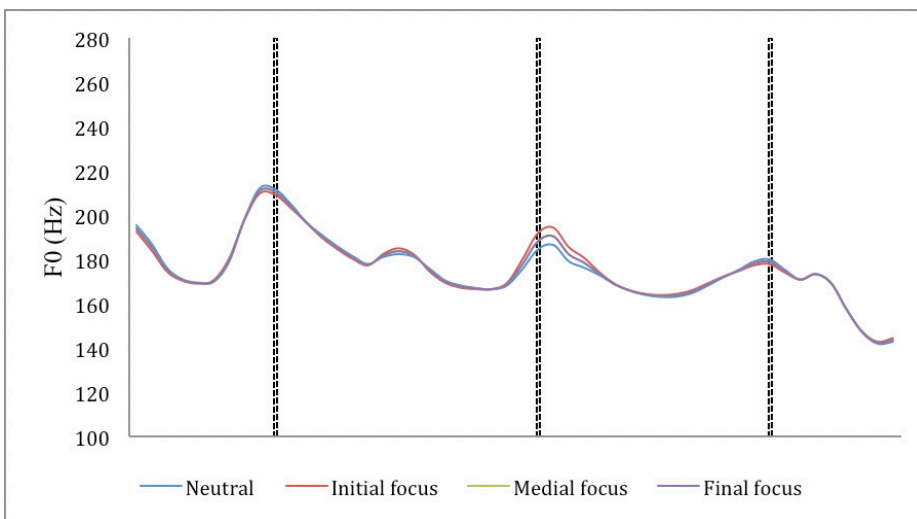


Figure 2. Time-normalised mean  $F_0$  contours of the averaged values across four unique sentences. Each curve in the figure represents each focus condition with an average of 12 repetitions across four unique sentences that account for 48 tokens in total.

further lowering for medial and final focus, Sentence 3 does not show lowering for medial focus whereas Sentence 4 only shows indication following the initial on-focus words. Overall, the time-normalised mean  $F_0$  contours at hand have provided the study with mixed and ambiguous results. Whereas the contour of mean  $F_0$  averaged across unique sentences shows no trace of PFC, isolated sentences extracted from the averaged results have supplied the study with inconsistent hints of PFC the language. At this juncture, it is safe to say that PFC is seemingly absent in the language; however, support from the statistical analysis will reveal further insights.

The fundamental approach for the analysis is to compare the values between

the on-focus and non-focused words. In order to make comparisons between on-focus and non-focused conditions in three different locations, four prosodic parameters namely max  $F_0$ , mean  $F_0$ , intensity and duration were regarded as dependent variables while focus (focus, non-focused/neutral) and focus locations (initial, medial, final) were taken as independent variables. The study also compared the post-focus words in both conditions by examining the areas following the initial and medial focus locations. The final focus location was not included in the analysis of post-focused words because there was no word following the fifth word. This study analysed the dependent variables in two ways by averaging the values across unique

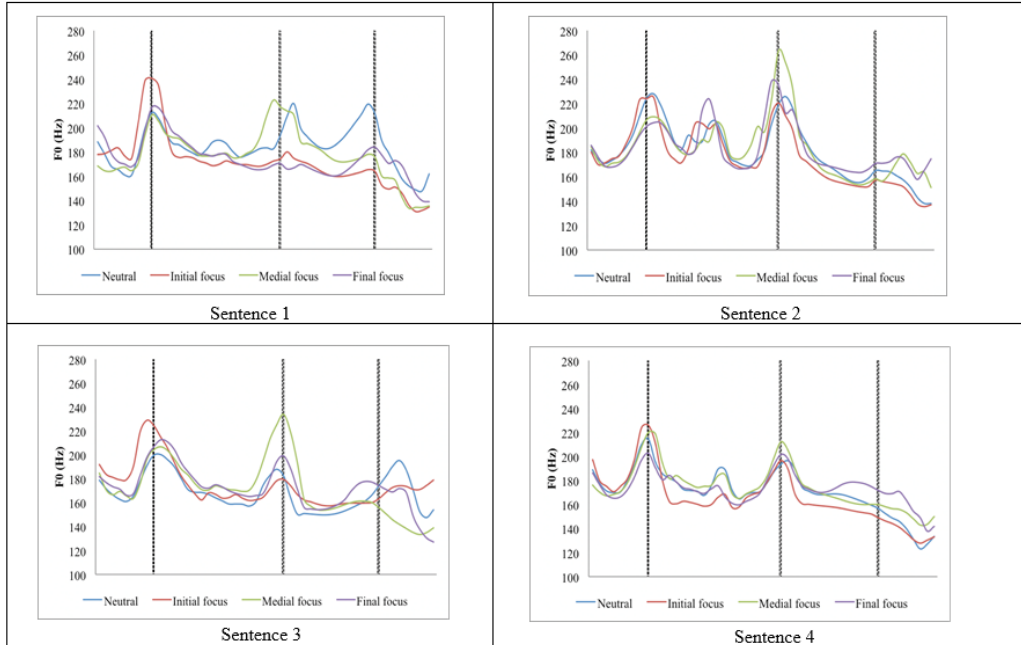


Figure 3. Time-normalised mean  $F_0$  contours of the values across four unique sentences with each curve representing an average of 12 repetitions by 12 participants.

sentences and by treating the four unique sentences as a within-subject factor. For the averaged values across unique sentences, 2-way repeated-measures ANOVAs were employed while 3-way ANOVAs were applied to analyse the effects on four unique sentences in isolation.

The results from the analysis of 2-way ANOVAs showed that for the on-focused words, max  $F_0$ , intensity, and duration are significantly higher than their non-focused counterparts. However, there is no significant difference for mean  $F_0$  between the on-focused and non-focused words. The effect of focus locations is significant for all parameters except duration. In the focused words, all parameters have higher values as compared to the later location: max  $F_0$ : 246.612 vs. 206.252, mean  $F_0$ : 194.693 vs.

165.263, intensity: 72.037 vs. 68.466, and duration: 402.303 vs 395.263. There is only one significant interaction between focus and focus location which is for max  $F_0$ :  $F(2, 22) = 15.931, p < 0.001$ .

The comparison between post-focused words in on-focused and non-focused conditions shows no evidence of significant difference in all four prosodic parameters (see Table 2). Meanwhile, the effect of focused location displays significant difference in mean  $F_0$  ( $F [1, 11] = 8.033, p = 0.016$ ), intensity ( $F [1, 11] = 24.795, p < 0.001$ ) and duration ( $F [1, 11] = 696.319, p < 0.001$ ). There is only one significant interaction of focus and focus location found which is indicated in duration ( $F [1, 11] = 8.175, p = 0.016$ ).

Table 2

*Results of 2-way repeated-measures ANOVAs for all prosodic parameters on their averaged values across four unique sentences*

On-focus words			
	Focus (df = 1, 11)	Location (df = 2, 22)	Focus * Location (df = 2, 22)
Max $F_0$	F = 26.571 <b>p = 0.000*</b>	F = 6.880 <b>p = 0.005*</b>	F = 15.931 <b>p = 0.000*</b>
Mean $F_0$	F = 4.107 p = 0.068	F = 15.655 <b>p = 0.000*</b>	F = 1.861 p = 0.179
Intensity	F = 12.842 <b>p = 0.004*</b>	F = 37.528 <b>p = 0.000*</b>	F = 3.429 p = 0.051
Duration	F = 14.398 <b>p = 0.003*</b>	F = 0.196 p = 0.750	F = 2.402 p = 0.114
Post-focus words			
	Focus (df = 1, 11)	Location (df = 1, 11)	Focus * Location (df = 1, 11)
Max $F_0$	F = 0.138 p = 0.717	F = 4.698 p = 0.053	F = 0.005 p = 0.944

Table 2 (Continued)

Post-focus words			
	Focus (df = 1, 11)	Location (df = 1, 11)	Focus * Location (df = 1, 11)
Mean F <sub>0</sub>	F = 0.264 p = 0.618	F = 8.033 <b>p = 0.016*</b>	F = 0.835 p = 0.381
Intensity	F = 0.059 p = 0.813	F = 24.795 <b>p = 0.000*</b>	F = 3.429 p = 0.092
Duration	F = 0.25 p = 0.876	F = 696.319 <b>p = 0.000*</b>	F = 8.175 p = 0.016*

On the other hand, comparisons using 3-way ANOVAs analysing four unique sentences as a within-subject factor revealed a similar fashion in the main effect of focus. As can be seen in Table 3, the main effect of focus shows a significant difference in max F<sub>0</sub>, intensity and duration parameters while mean F<sub>0</sub> shows no significant difference. The main effect of location is significant in max F<sub>0</sub>, mean F<sub>0</sub> and intensity whereas, for the effect of sentence, significant effects are found in max F<sub>0</sub>, intensity and duration. Note that the interpretation of the main effect of the sentence is not directly explainable considering the differences in structural make-up underlying each unique sentence. In addition, the interaction of sentence and focus is only evident in intensity such that  $F(3, 33) = 4.227, p = 0.12$ .

The analysis of post-focused words on the basis of unique sentences revealed no significant difference in the main effect of focus in three prosodic parameters which max F<sub>0</sub>, mean F<sub>0</sub>, and intensity. Instead, duration of post-focused words shows significant change ( $F[1, 11] = 10.814, p = 0.007$ ). Moreover, significant main effects of the sentence are observed in intensity and duration, while the main effect of location is found significant in mean F<sub>0</sub>, intensity, and duration. Lastly, the interaction between sentence and focus indicates moderate significant effect for max F<sub>0</sub>, whereas intensity and duration illustrates high significant differences at  $F(3, 33) = 14.480, p < 0.001$  and  $F(3, 33) = 25.898, p < 0.001$ .

Table 3

Results of 3-way repeated-measures ANOVAs for all prosodic parameters on their values across four unique sentences

On-focus words				
	Focus (df=1, 11)	Location (df = 2, 22)	Sentence (df = 3, 33)	Sentence*Focus (df = 3, 33)
Max F <sub>0</sub>	F = 5.264 <b>p = 0.042*</b>	F = 11.419 <b>p = 0.000*</b>	F = 4.140 <b>p = 0.014*</b>	F = 1.312 p = 0.287

Table 3 (Continued)

On-focus words				
	Focus (df=1, 11)	Location (df = 2, 22)	Sentence (df = 3, 33)	Sentence*Focus (df = 3, 33)
Mean F <sub>0</sub>	F = 4.107 p = 0.068	F = 15.655 <b>p = 0.000*</b>	F = 2.855 p = 0.052	F = 0.587 p = 0.628
Intensity	F = 12.842 <b>p = 0.004*</b>	F = 37.528 <b>p = 0.000*</b>	F = 3.308 <b>p = 0.032*</b>	F = 4.227 <b>p = 0.012*</b>
Duration	F = 14.398 <b>p = 0.003*</b>	F = 0.196 p = 0.824	F = 23.738 <b>p = 0.000*</b>	F = 0.832 p = 0.486
Post-focus words				
	Focus (df = 1, 11)	Location (df = 1, 11)	Sentence (df = 3, 33)	Sentence* Focus (df = 3, 33)
Max F <sub>0</sub>	F = 3.007 p = 0.111	F = 4.014 p = 0.07	F = 1.094 p = 0.365	F = 2.947 p = <b>0.047*</b>
Mean F <sub>0</sub>	F = 1.274 p = 0.283	F = 6.314 <b>p = 0.029*</b>	F = 0.758 p = 0.394	F = 1.752 p = 0.176
Intensity	F = 4.282 p = 0.063	F = 9.882 <b>p = 0.000*</b>	F = 14.348 <b>p = 0.003*</b>	F = 14.480 <b>p = 0.000*</b>
Duration	F = 10.814 <b>p = 0.007*</b>	F = 25.898 <b>p = 0.000*</b>	F = 3.641 <b>p = 0.023*</b>	F = 25.898 p = 0.000*

Notes: Statistically significant differences (ANOVA) are indicated by \* $p < 0.05$ ; N = 12; df = degrees of freedom

## Discussions

The objective of the study was to describe the prosodic realisation of focus in Malay. To achieve the objective, two questions were raised as to whether there is a difference in the variations of F<sub>0</sub> contour, intensity, and duration between focused and non-focused conditions. The study also aims to search for a possibility of PFC in the language by comparing the focused and non-focused correlates at the region after the on-focus words.

Table 2 and 3 from the previous section exemplify clear differences of the implicated

prosodic parameters due to the effect of focus in Malay. According to the current findings, the words under focus appear to have a higher pitch, higher intensity, and longer duration although pitch-wise, the parameter of mean F<sub>0</sub> indicates no substantial difference under this effect. In terms of focus locations, similar evidence of a heightened pitch for the focused words is found in Hong Kong Cantonese, with the initial location showing a more pronounced contour than the others in the sentence (Wang et al., 2011). Indeed, the fact that the initial focus shows the highest pitch

contour as compared to other focus locations tallies with the observations implicated on different effects of focus in different locations (Cooper et al., 1985). These results are well supported looking at the consistent display of significance across two levels of analysis employed in the study which are comparisons across values of unique sentences and across the averaged values of these unique sentences. These findings have directly addressed the first question raised in this study and they are found to be in accord with other studies on prosodic realisation of focus in this regard (Cooper et al., 1985; Eady & Cooper, 1986; Féry & Kugler, 2008; Heldner, 2003; Xu & Xu, 2005).

As for the post-focused words, the results from the analyses suggest a series of salient outcomes such that there is an absence of significant lowering or compression in all prosodic parameters: pitch, intensity, and duration. Again, these prosodic parameters manifest a consistent pattern on both levels of analysis, with the exception of duration that shows significant difference across unique sentences. Put together, these findings fit very well with the three-zone pitch range adjustments model as suggested in the earlier section (Liu & Xu, 2005; Xu, 1999). These findings confirmed the two zones of pitch adjustments from the model, namely the expansion under focus and the consequential compression following it as a result of focus. A follow-up study will further illustrate the fit of the intonational model in the pre-focus regions in Malay. In such a case, findings will illustrate if Malay shows a presence/absence

of a neutral shape in the pre-focus words with minimal variation of  $F_0$  range.

Referring to the second question presented in the earlier part of this study, Malay appears to exhibit no evidence of pitch and intensity lowering on the post-focus words, thus implying that the language belongs to the group without PFC. The literature review has indicated several languages that are characterised as having no PFC such as Yi, Deang, and Wa (Wang et al., 2011) and comparisons between these languages have shown a similar pattern of  $F_0$  with that of Malay. Malay also seems to share an expanse of similarities with Taiwanese such that the raising of  $F_0$ , intensity and duration appear to be the major acoustic correlates for signalling focus (Chen et al., 2009). Moreover, a comparison of the results with another language that is devoid of PFC noted a similar fashion in terms of time-normalised mean pitch contour. Specifically in Hong Kong Cantonese, post-focus words display the only negligible observable differences in pitch contours between focused and non-focused contours, with a rather minimal focus-related variability (Wu & Xu, 2010). Closer scrutiny further confirms the mutual similarities that are shared between the two languages through the indication of non-significant mean  $F_0$  under the effect of focus. At this juncture, this study has arrived at several deductions. First, Malay applies the same focus signalling strategy as the other languages without PFC by utilising duration, intensity and sometimes pitch parameters in an utterance. This is contrary to the PFC



languages which showed the utilisation of the lowering of prosodic parameters following the on-focus words. Secondly, this study has led to a strong affirmation in the typological categorisation of languages on the basis of PFC. A series of studies have substantiated a clear-cut distinction between languages positing PFC and the ones without by showing no traits of ambiguous overlaps in the prosodic measures. These findings have demonstrated the notion of PFC as a strong and robust element to designate languages into different groups that can probably be explained by a certain ancestral origin.

This combination of findings now leads the discussion to the subsequent issue of language typological differences. Notably, this study was carried out with the purpose of addressing the bigger hypothesis of linguistic typology as proposed by Xu (2011). The highly endorsed proposal suggests a proto-language called Nostratic superfamily to be the common origin for the grouping of all modern PFC languages, within which lies seven different language families which are Indo-European, Uralic, Altaic, Afroasiatic, Dravidian, Kartvelian, and Eskimo-Aleut. Malay, however, is not hypothesised to be in the same grouping since Austronesian origin does not constitute the grouping of Nostratic superfamily. Hence, findings would be informative in testing the viability of the hypothesis by confirming the exclusion of the Austronesian family from the hypothesis. Indeed, this is the case with the current findings; however, a note of caution is due since a negative

proof could leave to a logical fallacy. It could be a dangerous attempt to base the soundness of the hypothesis upon the lack of evidence from other language families that are not implicated. Xu (2011) noted that languages from the Dravidian, Kartvelian, and Eskimo-Aleut families had not yet been tested for PFC, thus future research for these languages should be carried out. In summary, the results have corroborated supplementary support for the hypothesis involving the ancestral origin of modern PFC languages.

## CONCLUSION

The present study concurs that the prosodic realisations of focus in Malay were better described within the articulatory-functional framework. It has provided an alternative outlook to the prosodic inventory of Malay looking at the complications that emerged from the adaptation of the A-M framework of intonation. Under the effect of focus, Malay illustrates heightened pitch, increased intensity, and elongated duration; a series of characteristics that have been generally implied in the literature across languages. The language also works in agreement with the three-zone model of pitch range adjustments consisting of a neutral shape and intact structure for the pre-focus words, an expansion under focus, and lowering for the words following the on-focus words. Notably, the present study had unearthed a novel feature of Malay which was not previously discovered, namely the absence of post-focused compression (PFC). This adds another layer of description to the

language such that it exhibits similarities with several existing languages on the Southern coast of China and Taiwan. More importantly, this renders support that is central to the hypothesised origin of the language, which belongs to the Austronesian language family.

The authors recognize several improvements for a better representation of the study. Essentially, the study could have been better described with rectifications on the overall experimental design. Better considerations on the choice of words and the method of repetitions of sentences would be necessary for the ease of segmentation and to avoid the collection of anomalous and random data. The authors noted difficulties in segmenting the words from the raw data, which might have led to less accurate markings of word boundaries. Although word onsets were controlled by sonorants, it was not fulfilled on all occasions. Moreover, the study collected repetitions of different focus locations across unique target sentences rather than across identical sentences. The method of repetitions could have been better implemented if every target sentence consisting of different focus locations was randomised and repeated during the recording session. Further study will also reveal the feature of pre-focused words in Malay under the effect of focus since this study did not approach the query in that regard. Lastly, a perception experiment could be supplemented to the production experiment to seek for the effectiveness of focus signalling strategy underlying the language. Since the language is devoid

of PFC, which is an effective element for focus recognition, further study will reveal other focus elicitation approach that could be employed in the language.

## ACKNOWLEDGEMENT

The authors thank the UCL Division of Psychology and Language Sciences for their support in terms of financing the project (Project number: SHaPS-2014-YX-013) and the facilities provided. A special thanks to all respondents for their support and cooperation throughout data collection.

## REFERENCES

- Anceaux, J. (1965). Linguistic theories about the Austronesian homeland. *Bijdragen Tot De Taal-, Land- En Volkenkunde*, 121(4), 417-432.
- Bellwood, P. (1997). *Prehistory of the Indo-Malaysian Archipelago*. Honolulu: University of Hawaii.
- Boersma, P., & Weenink, D. (2015). *Praat: Doing phonetics by computer [Computer program]*. Retrieved August 1, 2015, from <http://www.praat.org/>
- Bolinger, D. (1958). Intonation: Levels versus configurations. In Bolinger (Ed.), *Forms of English: Accent, morpheme, order* (pp. 17-56). Cambridge: Harvard University Press.
- Botinis, A., Fourakis, M., & Gawronska, B. (1999). Focus identification in English, Greek and Swedish. In *Proceedings of the 14th International Congress of Phonetic Sciences* (pp. 1557-1560). San Francisco: University of California, Berkeley.
- Bruce, G. (1982). Developing the Swedish intonation model. *Lund University, Department of Linguistics Working Papers*, 22(1982), 51-116.

- Chahal, D. (2003). Phonetic cues to prominence levels in Lebanese Arabic. In *Proceedings of the Fifteenth International Congress of Phonetic Sciences* (pp. 2067-2070). Barcelona: ICPhS Organizing Committee.
- Chen, S.-W., Wang, B., & Xu, Y. (2009). Closely related languages, different ways of realizing focus. In *Proceedings of Interspeech* (pp. 1007-1010). Brighton: Causal Productions.
- Cooper, W. E., Eady, S. J., & Mueller, P. R. (1985). Acoustical aspects of contrastive stress in question-answer contexts. *The Journal of the Acoustical Society of America*, 77(6), 2142-2156.
- Dohen, M., & Lævenbruck, H. (2004). Pre-focal rephrasing, focal-enhancement and post-focal deaccentuation in French. In *Proceedings of the 8th International Conference on Spoken Language Processing* (1313-1316). Jeju, Korea: Sunjin Printing Company.
- Eady, S. J., & Cooper, W. E. (1986). Speech intonation and focus location in matched statements and questions. *The Journal of the Acoustical Society of America*, 80(2), 402-415.
- Féry, C. (2001). *Focus and phrasing in French*. Berlin: Akademie-Verlag.
- Féry, C. & Kugler, F. (2008). Pitch accent scaling on given, new and focused constituents in German. *Journal of Phonetics*, 36(4), 680-703.
- Heldner, M. (2003). On the reliability of overall intensity and spectral emphasis as acoustic correlates of focal accents in Swedish. *Journal of Phonetics*, 31(1), 39-62.
- Hellmuth, S. (2006). Focus-related pitch range manipulation (and peak alignment effects) in Egyptian Arabic. In *Proceedings of Speech Prosody 2006* (PS4-12-164). Dresden, Germany: TUD Press.
- Howie, J. M. (1974). On the domain of tone in Mandarin. *Phonetica*, 30(3), 129-148.
- IBM Corp. (2013). *IBM SPSS Statistics for Mac, Version 22.0*. Armonk: IBM Corp.
- Ipek, C. (2011). Phonetic realization of focus with no on-focus pitch range expansion in Turkish. In *Proceedings of the 17th International Congress of Phonetic Sciences Hong Kong* (pp. 140-143). Hong Kong: City University of Hong Kong.
- Ishihara, S. (2002). Syntax-phonology interface of Wh- Constructions in Japanese. In *Proceedings of Tokyo Conference on Psycholinguistics (TCP 2002)* (pp. 165-189). Tokyo: Hituzi Syobo.
- Jin, S. (1996). *An acoustic study of sentence stress in Mandarin Chinese*. Retrieved September 30, 2015, from <https://etd.ohiolink.edu/>
- Jun, S. A., & Lee, H. J. (1998). Phonetic and phonological markers of contrastive focus in Korean. *Proceedings of the 5th International Conference on Spoken Language Processing*, 4(1), 1295-1298.
- Kügler, F., & Skopeteas, S. (2007). On the universality of prosodic reflexes of contrast: The case of Yucatec Maya. In *Proceedings of the 16th International Congress of Phonetic Sciences* (pp. 1025-1028). Saarbrücken, Germany: TAL.
- Lee, Y. C., & Xu, Y. (2010). Phonetic realization of contrastive focus in Korean. In *Proceedings of Speech Prosody 2010* (pp. 1-4). Chicago: International Speech Communication Association.
- Liu, F., & Xu, Y. (2005). Parallel encoding of focus and interrogative meaning in Mandarin intonation. *Phonetica*, 62(2-4), 70-87.
- Ladd, D. R. (2008). *Intonational phonology* (2nd ed.). Cambridge: Cambridge University Press.
- Lehiste, I. (1975). Suprasegmentals. *Language*, 51(3), 736-740. <http://doi.org/10.1371/journal.pone.0005772>

- Lehiste, I., & Peterson, G. E. (1961). Some basic considerations in the analysis of intonation. *The Journal of the Acoustical Society of America*, 33(4), 419-425.
- Liu, F., & Xu, Y. (2005). Parallel encoding of focus and interrogative meaning in mandarin intonation. *Phonetica*, 62(2-4), 70-87.
- Mixdorff, H. (2004). Quantitative tone and intonation modelling across languages. In *Proceedings of International Symposium on Tonal Aspects of Languages: With Emphasis on Tone Languages* (pp. 137-142). Beijing: The Institute of Linguistics in Chinese Academy of Social Sciences.
- Patil, U., Kentner, G., Gollrad, A., Kügler, F., Féry, C., & Vasishth, S. (2008). Focus, word order and intonation in Hindi. *Journal of South Asian Linguistics*, 1(1), 55-72.
- Rialland, A., & Robert, S. (2001). The intonational system of Wolof. *Linguistics*, 39(5), 893-939.
- Rump, H. H., & Collier, R. (1996). Focus conditions and the prominence of pitch-accented syllables. *Language and Speech*, 39(1), 1-17.
- Taheri, A. M., & Xu, Y. (2012). Phonetic realization of prosodic focus in Persian. *Proceedings of Speech and Prosody 2012* (pp. 326-329). Shanghai: Tongji University Press.
- van Geldern, R. H. (1966). A note on relations between the art styles of the Maori and Ancient China. In F. B. Verlag & Söhne (Eds.), *Wiener Beiträge zur Kulturgeschichte und Linguistik* (pp. 45-86). Band XV, Vienna: Verlag Ferdinand Berger & Söhne Oho.
- Wang, B., Wang, L., & Kadir, T. (2011). Prosodic encoding of focus in six languages in China. In *Proceedings of the 17th International Congress of Phonetic Sciences* (pp. 144-147). Hong Kong: City University.
- Wolters, O. W. (1982). *History, culture, and region in Southeast Asian perspectives*. Singapore: Institute of Southeast Asian Studies.
- Wu, W. L., & Xu, Y. (2010). Prosodic focus in Hong Kong Cantonese without post-focus compression. *Proceedings of Speech Prosody Chicago 2010*, 100040(1), 1-4.
- Xu, Y. (1999). Effects of tone and focus on the formation and alignment of F<sub>0</sub> contours. *Journal of Phonetics*, 27(1), 55-105.
- Xu, Y. (2011). *Post-focus compression: Cross-linguistic distribution and historical origin*. Retrieved September 30, 2015, from <http://discovery.ucl.ac.uk/1322251/>
- Xu, Y. (2013). ProsodyPro - A tool for large-scale systematic prosody analysis. In *Proceedings of Tools and Resources for the Analysis of Speech Prosody (TRASP 2013)* (pp. 7-10.). Aix-en-Provence, France: Laboratoire Parole et Langue.
- Xu, Y., Chen, S., & Wang, B. (2012). Prosodic focus with and without a post-focus compression (PFC): A typological divide within the same language family? *The Linguistic Review*, 29(1), 131-147.
- Xu, Y., Lee, A., Prom-on, S., & Liu, F. (2015). Explaining the PENTA model: A reply to Arvaniti and Ladd. *Phonology*, 32(3), 505-535.
- Xu, Y., & Xu, C. (2005). Phonetic realization of focus in English declarative intonation. *Journal of Phonetics*, 33(2), 159-197.
- Zerbian, S., Genzel, S., & Kügler, F. (2010). Experimental work on prosodically-marked information structure in selected African languages (Afroasiatic and Niger-Congo). In *Proceedings of Speech Prosody Chicago* (pp. 1-4). Illinois: Illinois University Press.