Interlanguage Influence in Cues of Narrow Focus: a study of Hong Kong English

Richard Yohann Gananathan, Yanjun Yin, Pik Ki Peggy Mok

The Chinese University of Hong Kong
rgan@cuhk.edu.hk, yj.yin@cuhk.edu.hk, peggymok@cuhk.edu.hk

ABSTRACT

The acoustic properties of narrow focus are language-specific. Post-focus compression of pitch is one of the most prominent factors characterizing focus in English, but it is not present in Cantonese, where duration and intensity are more dominant. Hong Kong English (HKE) is an emerging variety of English in Hong Kong, characterized by speakers who grow up learning English from a young age, but with input highly influenced from the dominant language Cantonese.

Experiments were conducted testing production and perception of focus in HKE speakers. An overall tendency was observed for speakers to have accurate perception, but production that was non-characteristic of traditional varieties of English. Analysis of the highest-scoring participants found that their usage of pitch to mark focus was minimal, but usage of duration was significant.

Keywords: Hong Kong English, focus, perception, production, duration

1. INTRODUCTION

Focus is a technique used by speakers to emphasize and point out new material in a sentence [4]. Its acoustic correlates in English have widely been shown to involve pitch fluctuations, as well as intensity and duration effects [8]. More recent studies have shown that in addition to the above effects on the focus peak, non-focused elements are also affected. After the on-focus peak, post-focal elements were found to have lowered and narrowed pitch, in a phenomenon known as post-focus compression [8, 9]. In Cantonese, studies have shown results demonstrating the importance of intensity and duration effects [2, 3, 5, 10]. While some [5] suggest that on-focus pitch expansion may be exhibited, this has been argued against, with support from a more comprehensive study [10]. Post-focus compression however, does not seem to occur in Cantonese [10].

Current research on Hong Kong English (HKE) focus is scarce. Fung & Mok [1] have done a production study that found that HKE focus was characterized by what they claim to be a mix of English and Cantonese cues. It was found to have f0 pitch range expansion like English, yet lacked post-focus compression. However, their study did not use the usual method of using questions to elicit focus data from HKE speakers, but rather gave the focus information to the participants directly by boldfacing the target focus words. They claimed this experimental procedure was necessary in order to elicit focus from HKE speakers, as speakers did not seem to apply focus in question response style experiments. While their findings go part of the way towards describing what HKE focus is like, the data may not be very natural, and it is unclear how HKE speakers would realize focus under question-response setting. In order to investigate focus realization in HKE, further study is needed, which the experiments described in this paper seek to contribute to.

This research consists of two studies, a production experiment and a perception experiment. Both experiments used questions to elicit focus. We found that while the HKE speakers' responses to the questions designed to elicit focus were rarely understood to have the intended focus, the HKE speakers were very proficient at perceiving focus as spoken by a native British English speaker. The production experiment results also show that while many of the low-performing HKE speakers did not appear to acoustically mark focus at all in their responses, for some of the high-performing HKE speakers, duration plays a more important role than pitch, which is often considered to be the primary acoustic property of focus in traditional varieties of English.

2. METHODOLOGY

Two experiments were conducted, a production experiment and a perception experiment. A total of 42 HKE speakers participated in the experiment. The participants were all university students from the Chinese University of Hong Kong with Cantonese as a first language and HKE as a second. Each participant completed both the production and perception experiments. Three of the participants were excluded on the basis of two being found to have native languages other than Cantonese, and the last being found to have hearing loss. In addition to the HKE participants, one native British English speaker was also recorded as a control, although she

did not participate in the experiment directly but simply read out the intended stimuli.

10 unique sentences were used in the experiment. 4 for the production experiment, 4 for the perception experiment, and 2 as practice sentences that were not analysed. In each set of 4 sentences, for half of the sentences, only the initial, medial, and final positions were compared, while for the other half, the pre-medial, medial, and post-medial positions were compared. This was done in order to observe more phenomena.

2.1. Production experiment

The materials were presented to participants through PowerPoint slides. The first few slides provide written instructions to the participants, followed by two trial tokens, of the same style as real tokens, which habituate the participant to the format of the experiment. After the trial run, the participants confirm that they are ready to start the experiment, and the experiment begins.

Each slide initially presents the participant with a written sentence. Participants are given two seconds to read this sentence, and then a recording of the question is played to the participant. The recorded question would repeat the content of the written sentence and ask if it is true, but with one word changed. The participants were instructed to respond to this question verbally, in the form that uses the full sentence displayed earlier. This response was recorded, and the participant would click to proceed to the next token. This design is intended to suggest the participant repeat the original sentence with focus on the word that is different from the visually presented sentence.

For example: a sentence like “The weather in Britain is often very breezy.” would be displayed on a slide, and then a recording asking “Is the weather in Sweden often very breezy?” would be played. The participant would then respond “No, the weather in Britain is often very breezy.” This example would elicit focus on “Britain”.

A total of 12 tokens were recorded per speaker, consisting of 4 different sentences, each carrying initial/pre-medial focus, medial focus, and final/post-medial focus conditions. The order of the tokens was randomized, but the order presented to each participant was identical.

The recordings of the participant responses were then listened to by five judges, two of whom were native speakers of Canadian English, and the other three of whom were non-native English speakers. All had phonetic training. The five judges selected the word they believed the speaker was trying to focus, and gave a rating out of 7 of how certain they were of that judgment.

In addition to the judgements, the sound files for the sentences produced by 8 speakers were segmented and further analyzed acoustically. These 8 speakers were selected on the basis of their scores in the experiment. 3 of them had high scores in both the production and perception experiments (See section 2.2 for the perception experiment), 3 had high scores in the perception experiment but low scores in the production experiment, and the last 2 had low scores in both the production and the perception experiments. All of the sonorants were segmented and the Praat script Prosody Pro was used to extract the data [6].

2.2. Perception experiment

The perception experiment was also presented to participants using PowerPoint. Again, two trial tokens and a confirmation were put in before the experiment begins.

In each slide, a written question appears alone first. After two seconds, possible answers A., B., and C. are played aloud, and the participant is instructed to write down the answer they judge to be correct on a piece of paper provided. The participant is given the option to re-play any of the possible answers before writing his answer and choosing to move on to the next token.

The answers to the written question are identical to the question, but with one word changed. The same word is replaced in answers A., B., and C., but the focus is on a different word in each answer. This tests the participants' perception by seeing if they can pick the answer with the correct focus, that is, where focus is on the changed word.

For example: a question such as “Should someone older than Michael's sister purchase the tickets?” would be displayed on the slide, and three answers with the format “Someone older than Michael's brother should purchase the tickets.” will be played. In A., the focus is on “purchase”, in B., the focus is on “older”, and in C., the focus is on “brother” - the correct answer.

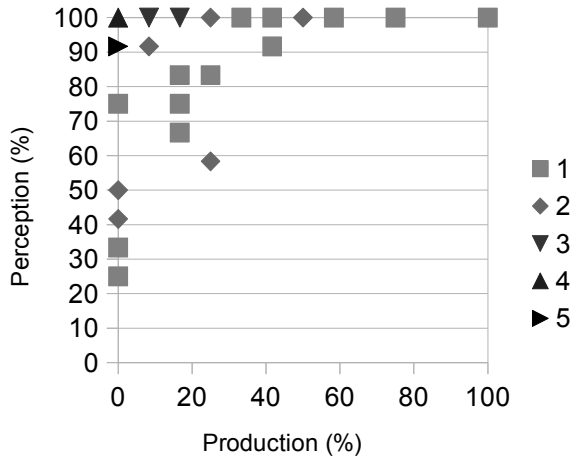
The format of the trials was similar to the tokens collected in the previous experiment with 12 total trials per speaker. The trials consisted of 4 different sentences repeated with correct focus in each condition (one of initial/pre-medial, medial, and final/post-medial). The collected data was then gathered and is presented in the next section.

3. RESULTS

Figure 1 shows that while individual abilities varied, the HKE speakers in general had more accurate

performance on the perception task than the production task.

Figure 1: A scatter plot of each speaker's perception × production scores from the two experiments. The legend indicates the number of points with the same score for both perception and production.



The perception scores were calculated by taking the percentage of questions in the perception experiment in which they correctly judged the intended focused element. The production scores were calculated by taking the percentage of their responses the five judges were able to correctly judge the intended focused element for. (The judge[s' confidence rating was not considered.) The judges overall did not have a very high rate of consistency, as often times they had to simply make guesses; however, for 21% of the tokens, all five of the judges gave identical word judgements. The judges likely were more consistent in judging the more proficient HKE participant-speakers.

Figure 2 and Figure 3 compare the pitch contours for Sentence 6 of the British English speaker and the one HKE speaker with a 100 percent score for both production and perception. Time is normalized at a rate where 10 points were measured per sonorant. This allows comparison between speakers, as each normalized time point refers to the same part of the sentence. This specific HKE speaker was chosen for further study because she clearly demonstrated some use of pitch to mark focus. All of the other HKE speakers analysed did not use pitch to mark focus very clearly. It should also be noted that this comparison is vulnerable to some variation between the two speakers from the difference in circumstances under which the data was obtained; the British English speaker was recorded reading lines, while the HKE speaker responded to questions.

Figure 2: The HKE speaker's pitch contours.

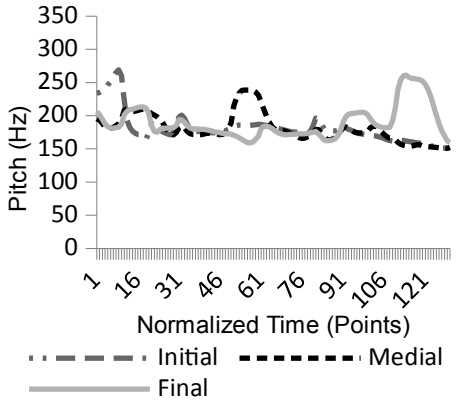
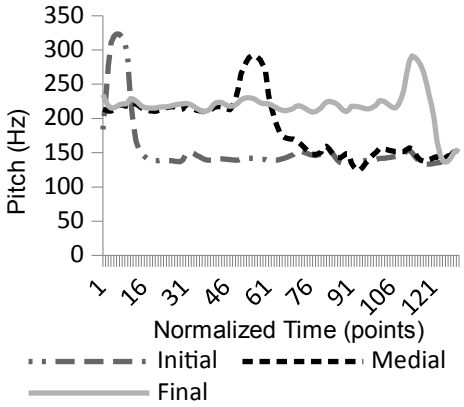


Figure 3: The British English speaker's pitch contours.



Here we notice that although the HKE speaker used pitch to mark on-focus raising, she did not exhibit post-focus compression as the British English speaker did.

While pitch was only found to play a significant role in the aforementioned one speaker, some of the other speakers seem to be using different cues, namely duration, to a greater extent than pitch. Figures 4 and 5 show the relation between the duration of an element and the position of focus

Note that in these graphs, “1st” and ”3rd” refers to either initial or pre-medial and final or post-medial focus, depending on which one was recorded for the sentence in question. (2 sentences recorded only initial/final, and the other 2 only recorded pre-medial and post-medial.) “2nd” always refers to the medial focus. For example: the bar labelled “3rd” in the X-axis's “Initial” group refers to the duration measurement taken at the initial position when the sentence had focus on the final (“3rd”) position; it is clearly shorter than the “1st” bar in the same group, which means it is shorter than the duration measurement taken at the initial (“1st”) position taken when the sentence had focus on the initial position. Also note that the data in the “Medial” group's bars is double that of the other four groups

as all 4 sentences had a medial position elicited, while the other 4 groups only include the 2 sentences that included the position in question.

A One-way ANOVA was conducted on each group of speakers to test significance. The 3 high-production high-perception participants produced significantly longer durations on on-focus elements than otherwise ($p < 0.000$). The the 3 high-perception low-production participants seemed to show such effects, but it was not as robust and is not statistically significant ($p = 0.337$). The low-perception low-production group's figures did not show any kind of correlation and are not shown here ($p = 0.517$). The data of the individual speakers within each group were also examined and found to be of an overall similar distribution, but statistical tests were not run because the data currently available is low.

Figure 4: The high-production high-perception group.

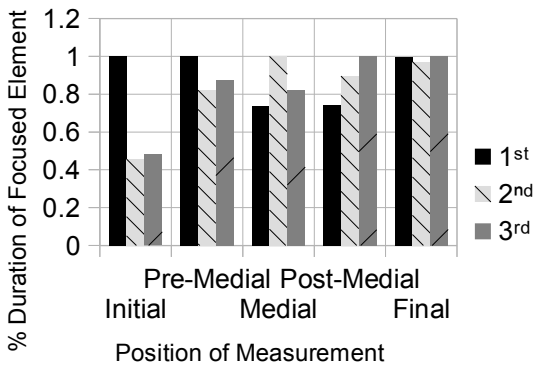
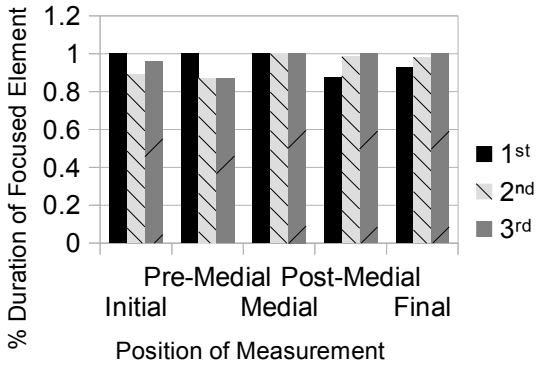


Figure 5: The high-perception low-production group.



4. DISCUSSION

As we can see from the results in Figure 1, the HKE speakers' responses to the focus-inducing questions were rarely judged to put focus on the intended position. This is despite the fact that many of the HKE speakers were proficient at perceiving focus in native British English speakers' utterances. This suggests that the HKE speakers in general may have better perception abilities than production abilities in terms of focus.

Looking further at the production data, only 1 speaker clearly used pitch in a meaningful way to indicate focus, but even she did not show any signs of post-focus compression. This is in accordance with the results from Fung & Mok's [1] study, which failed to find post-focus compressioneven though participants were explicitly instructed to focus selected words in a sentence. This study shows that a more naturalistic approach of using focus-inducing questions to elicit focus may still be a viable strategy when working with HKE speakers, and the acoustic cues indicating focus in HKE are simply different than in British English. Pitch is not the only acoustic cue for focus marking in English, and in fact duration plays a more important role in HKE. The other two speakers in the high-production high-perception group did not use pitch extensively to mark focus, but rather used duration. The judges successfully perceived this, as they were able to correctly judge the element intended to be focused.

The fact that the HKE speakers used duration more prominently than pitch is unsurprising, given that duration is a more important cue for marking focus than pitch in Cantonese as well [2, 3, 5, 10]. Post-focus compression in general is argued to be easily lost in language contact [7]. These interlanguage phonology effects provide a window on how new varieties of English are forming while being affected by a regional language.

4. CONCLUSION

The abilities of an individual speaker to produce and perceive focus distinctions are not necessarily the same. HKE speakers are found to generally have no difficulties perceiving focus as produced by a native English British speaker, but the HKE speakers' own production of focus is variable. For some of the less proficient speakers, it is unclear if there were any acoustic markings of focus in the sentences where we expected to see focus. For these speakers, English is likely more of an L2 language. Even for the most proficient speakers, however, none had post-focus compression as in dominant varieties of English, and few exhibited any on-focus pitch changes. Instead, increased duration was the primary factor they used to distinguish focus. Their productions of focus were therefore understood by the panel of judges to some degree.

This work is partially supported by the National Social Science Fund of China (10CYY009) and the Major Program for the National Social Science Fund of China (13&ZD189)

7. REFERENCES

[1] Fung, H. Mok, P. 2014. Realization of narrow focus in Hong Kong English declaratives: a pilot study. *Proc. 7th Int. Conf. Speech Prosody* Dublin, 964-968.

[2] Gu, W., Hirose, K. Fujisaki, H. 2006. The effect of paralinguistic emphasis on F0 contours of Cantonese speech. *Proc. 3rd Int. Conf. Speech Prosody* Dresden.

[3] Gu, W., Lee, T. 2007. Effects of tonal context and focus on Cantonese F0. *Proc. 16th ICPhS* Saarbrücken, 1033-1036.

[4] Halliday, M. A. K. 1967. Notes on transitivity and theme in English: part 2. *J. Linguistics* 3:2, 199-244.

[5] Man, V. 2002. Focus effects on Cantonese tones: an acoustic study. *Proc. 1st Int. Conf. Speech Prosody* Aix-en Provence, 467-470.

[6] Xu, Y. 2013. ProsodyPro: a tool for large-scale systematic prosody analysis. *Proc. Tools and Resources for the Analysis of Speech Prosody* Aix-en Provence, 7-10.

[7] Xu, Y., Chen, S.-W., Wang, B. 2012. Prosodic focus with and without post-focus compression: a typological divide within the same language family? *The Ling. Rev.* 29:131-147.

[8] Xu, Y. Xu, C. X. 2005. Phonetic realization of focus in English declarative intonation. *J. Phonetics* 33, 159-197.

[9] Xu, Y. Xu, C. X., Sun, X. 2004. On the temporal domain of focus. *Proc. 2nd Int. Conf. Speech Prosody* Nara, 81-84.

[10] Wu, W., Xu, Y. 2010. Prosodic focus in Hong Kong Cantonese without post-focus compression. *Proc. 5th Int. Conf. Speech Prosody* Chicago, 1-4.