



# Focus Effects on Cantonese Tones: An Acoustic Study

Vicky C. H. Man

Language Centre,  
Hong Kong Baptist University  
vickyman@hkbu.edu.hk

## Abstract

This paper discusses how the prosodic marking of sentential focus influences the acoustical pattern of F0 in Cantonese monosyllabic words. Results indicate that tone identities are maintained regardless of the focus conditions, and that a substantial pitch range expansion is associated with narrow focus. Pitch range is affected by the tone-focus interaction. The pitch range differences between the broad focus and narrow focus readings of each target word are insignificant.

## 1. Introduction

The notion of focus was put forward by Halliday in his early work on English intonation [1]. An individual word receiving the most prominence in an utterance is known as narrowly focused while broad focus is placed on a whole constituent (like a noun phrase) or on a whole sentence [2].

In a study of Mandarin sentence stress by Jin [3], narrow focus was associated with a substantial widening of pitch range. In particular, when the focused item was closer to the beginning of a sentence, its magnitude of expansion was greater than when it was located sentence-finally. The pitch range of the unstressed pre-focused words was not much affected, but that of the post-focused words was drastically reduced.

In Xu's investigation on how focus affects the formation of F0 contours in Mandarin [4], a similar pitch range phenomenon to Jin's study was observed. Xu suggested that there was a 'radical *asymmetry*' around the focus: the F0 range at the focus was substantially expanded; that after the focus was suppressed; and that before the focus did not deviated much from the non-focused condition.

In addition, Cantonese speakers often focus on certain words or phrases in natural conversations to sometimes make a contrast or a correction. It would therefore be interesting to find out how broad focus and narrow focus are manifested in Cantonese, whose tonal system is more complicated than that of Mandarin.

This study examines how focus influences the F0 curves of Cantonese monosyllabic words produced under the following conditions: citation with no focus, broad focus, and narrow focus signaled by contrastive stress. The extent to which focus affects the height and contour of Cantonese tones as well as the long-distance effect of narrow focus on the F0 of an unfocused word in the same sentence will also be explored.

## 2. Experimental Design

### 2.1. Sentence Materials

The production materials consisted of six declarative Cantonese sentences. Each contained two target words from

which the F0 measurements were taken. The carrier sentence was:

13 22 55 33 22                      22 21 22 33                      22  
k<sup>h</sup>øy wa: li: kɔ: hɛi \_\_\_\_\_ tsi: ɱ hɛi kɔ: \_\_\_\_\_ tsi:  
佢 話 呢 個 係                      字 唔 係 個                      字  
S/he says this CL is                      word not is CL                      word  
"S/he says this word is *target 1* not the word *target 2*."

Tonal environment was controlled, and possible tone sandhi and tone change effects were avoided by surrounding both target words with a Mid-Level (33) or a Mid-Low-Level (22) tone.

Prosodic comparison across languages is often difficult since language data used in previous studies are often not prosodically transcribed. To facilitate such a comparison and to resolve prosodic differences across languages, the Hallidayan model, a transcription method that allows a phonetic interpretation of rhythm and intonation of any language data, was adopted to assign the prosodic phrase boundaries in this study. The division into foot groups and intonation groups of the above carrier frame was determined auditorily after the production experiment, as illustrated in (2.1) below:

(2.1) ||<sub>A</sub> k<sup>h</sup>øy| wa: li: kɔ: hɛi|\_\_\_\_tsi:| |<sub>A</sub> |<sub>A</sub> ɱ| hɛi kɔ:|\_\_\_\_tsi:|  
"S/he says this word is | *target 1* | not the word | *target 2* ."

Two sets of six lexically contrastive words, each with the root morpheme of /si/ and /se/, were chosen for acoustical measurements. The /si/-set and /se/-set, each varied only in tone, were inserted in the sentence-medial and penultimate position of the carrier frame respectively, thus giving a corpus of six colloquial Cantonese sentences (Table 1).

1. k<sup>h</sup>øy wa: li: kɔ: hɛi **si: 55** 詩 tsi: ɱ hɛi kɔ: **se: 55** 睇 tsi:  
"S/he says this word is *poem* not the word *to lend*."
2. k<sup>h</sup>øy wa: li: kɔ: hɛi **si: 25** 史 tsi: ɱ hɛi kɔ: **se: 25** 寫 tsi:  
"S/he says this word is *history* not the word *to write*."
3. k<sup>h</sup>øy wa: li: kɔ: hɛi **si: 33** 嗜 tsi: ɱ hɛi kɔ: **se: 33** 瀉 tsi:  
"S/he says this word is *hobby* not the word *laxative*."
4. k<sup>h</sup>øy wa: li: kɔ: hɛi **si: 21** 時 tsi: ɱ hɛi kɔ: **se: 21** 蛇 tsi:  
"S/he says this word is *time* not the word *snake*."
5. k<sup>h</sup>øy wa: li: kɔ: hɛi **si: 23** 市 tsi: ɱ hɛi kɔ: **se: 23** 社 tsi:  
"S/he says this word is *market* not the word *society*."
6. k<sup>h</sup>øy wa: li: kɔ: hɛi **si: 22** 是 tsi: ɱ hɛi kɔ: **se: 22** 射 tsi:  
"S/he says this word is *yes* not the word *to shoot*."

Table 1: Test sentences.

The priming questions are listed in Table 2. In order to avoid any confusion with broad focus and to successfully elicit a non-focused production, subjects were first told to read the sentences in classroom repetition style.

Narrow focus on /si/ (Narrow-1):

k<sup>h</sup>øy hɛi-mɛi wa: li: kɔ: hɛi sy: tsi: m̩ hɛi kɔ: sɛ: tsi: a:  
 佢 係 咪 話 呢 個 係 書 字 唔 係 個 些 字 呀?  
 “Did s/he say this word is /sy/ not /sɛ/??”

Narrow Focus on /sɛ/ (Narrow-2)

k<sup>h</sup>øy hɛi-mɛi wa: li: kɔ: hɛi si: tsi: m̩ hɛi kɔ: sy: tsi: a:  
 佢 係 咪 話 呢 個 係 詩 字 唔 係 個 書 字 呀?  
 “Did s/he say this word is /si/ not /sy/??”

Broad focus:

nei kɔ:ŋ mət ye: a:  
 你 講 乜 野 呀?  
 “What did you say?”

Table 2: Priming questions.

The experiment consisted of 384 sentences: 6 tones, 4 focus patterns, 4 repetitions, 4 speakers (2 males, 2 females, all from Hong Kong). However, five sentences were read wrong and were discarded, resulting in a total of 379 sentences for acoustic analyses.

### 3. Results

#### 3.1. General Observation

F0 contours were normalized by extracting six F0 values at the 0%, 20%, 40%, 60%, 80%, and 100%-points of each tone contour. To prevent the higher F0 values of females from contributing too much to the mean values of each F0 contours, a log<sub>10</sub> conversion was performed on the F0 values obtained for each subject. This conversion accommodated the pitch range difference among speakers, especially that between males and females. The log values were then averaged across subjects. The mean log values were converted back to F0 values for display purposes [4].

Figures 1 to 4 show that the six contrastive lexical tones in Cantonese remain distinct regardless of focus conditions. The peak F0 of /si/ is found to be higher than that of /sɛ/. A lowering of F0 peak is particularly apparent when the High-Level (55) tone of /sɛ/ is preceded by a narrowly focused /si/ (Figure 2). Even when /sɛ/ is narrowly focused, its peak is not higher than that of /si/ (Figure 3). This lowering of F0 topline has been observed in non-tone languages such as English [5] and tone languages like Mandarin [3].

A comparison of Figure 1 with Figures 2-3 demonstrates that when the target words are narrowly focused, their pitch range is substantially wider than when they are produced in citation. When /si/ is compared with /sɛ/ in each focus condition, the pitch range of /sɛ/ is smaller than that of /si/ in citation, broad focus, and in the narrow focus reading of /si/. In particular, the pitch range of the narrowly focused /si/ is substantially expanded while that of the following non-focused /sɛ/ is severely lowered and compressed. This ‘asymmetry’ around focus has been reported in Mandarin [3], [4].

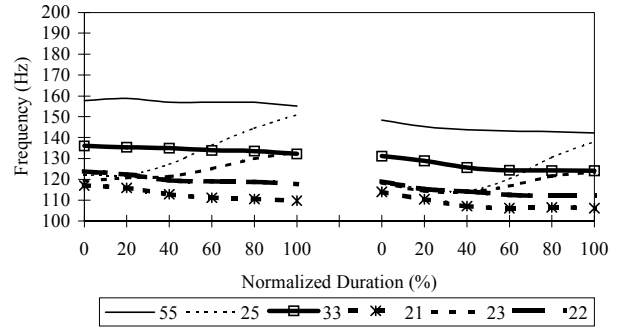


Figure 1: F0 contours of /si/ (left) and /sɛ/ (right) with respect to tones when produced in citation.

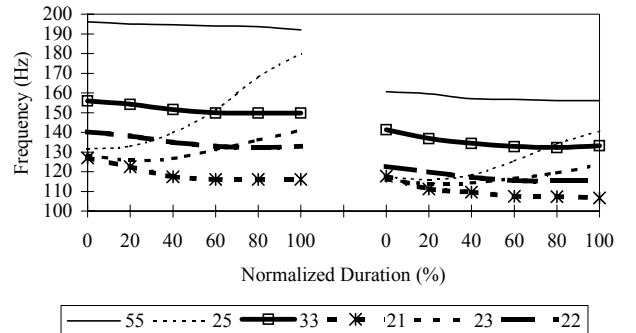


Figure 2: F0 contours of /si/ (left) and /sɛ/ (right) with respect to tones when /si/ is under narrow focus (Narrow-1).

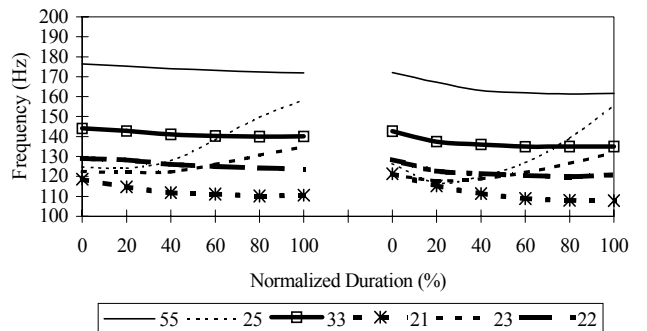


Figure 3: F0 contours of /si/ (left) and /sɛ/ (right) with respect to tones when /sɛ/ is under narrow focus (Narrow-2).

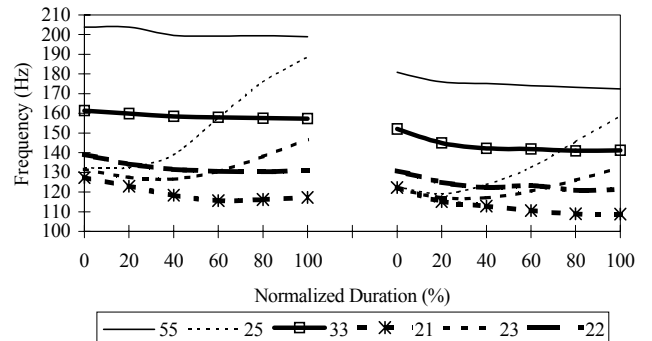


Figure 4: F0 contours of /si/ (left) and /sɛ/ (right) with respect to tones in broad focus.

### 3.2. Target Word /si/

In a two-way ANOVA analysis, a highly significant main effect of focus on pitch range ( $p=0.000$ ) is found. A post-hoc Scheffe multiple comparisons test indicates that the pitch range of /si/ in Narrow-1 is significantly greater than when /si/ is produced in citation ( $p=0.000$ ) and when it is unstressed before a focused /se/ in Narrow-2 ( $p=0.01$ ). When /si/ is in Narrow-2, its pitch range is not significantly wider than that in citation ( $p=0.651$ ). In broad focus, /si/ has a more expanded pitch range than when it is in citation ( $p=0.000$ ) or when it is followed by a contrastively stressed /se/ ( $p=0.000$ ). The pitch range of /si/ in broad focus, however, is not significantly greater than that in narrow focus (Narrow-1).

A two-way ANOVA also indicates that the effect of focus on the pitch range of /si/ is not the same for every lexical tone ( $p=0.000$ ). An examination of Figure 5 reveals that the High-Level (55), the Mid-Level (33), and the Mid-Low-Level (22) tones have a smaller pitch range than the High-Rising (25), Mid-Low-Rising (23) and the Mid-Low-Falling (21) tones. Results of an one-way ANOVA test reveal a highly significant difference in pitch range only for the High-Rising (25) tone ( $p=0.000$ ).

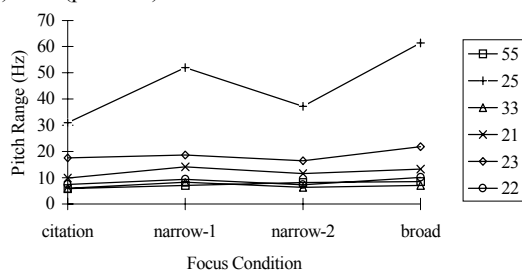


Figure 5: Mean pitch range values of /si/ as a function of tone and focus

### 3.3. Target Word /se/

Like /si/, the pitch range of /se/ is significantly related to focus ( $p=0.000$ ). The pitch range of /se/ in narrow focus (Narrow-2) is significantly greater than when /se/ is produced in citation ( $p=0.000$ ) or when it is preceded by a focused /si/ (Narrow-1) ( $p=0.001$ ). The same is true when /se/ is produced in broad focus. In broad focus, the pitch range of /se/ does not differ from that of /si/.

A two-way ANOVA indicates a significant interaction between tone and focus on the pitch range of /se/ ( $p=0.004$ ). In particular, a highly significant difference in pitch range is found for the High-Level (55) tone ( $p=0.000$ ) and the High-Rising (25) tone ( $p=0.002$ ). A post-hoc Scheffe multiple comparisons test shows that for the High-Level (55) tone, the pitch range of /se/ in narrow focus is significantly wider than when /se/ is in citation ( $p=0.000$ ) or when /se/ follows a narrowly focused /si/ (Narrow-1) ( $p=0.000$ ). For the High-Rising (25) tone, a significant pitch range difference is found between /se/ in narrow focus and citation ( $p=0.04$ ) as well as between /se/ in broad focus and citation ( $p=0.036$ ). A graphical illustration of the tone-focus interaction of /se/ is shown in Figure 6.

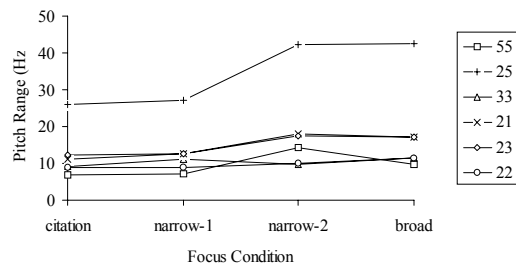


Figure 6: Mean pitch range values of /se/ as a function of tone and focus.

### 3.4. Comparison of /si/ and /se/

Two-tailed  $t$  tests suggest that the pitch range difference between /si/ and /se/ is not significant in citation ( $p=0.0391$ ) and in broad focus ( $p=0.08$ ). When /si/ is narrowly focused, its pitch range is more expanded than that of the following non-focused /se/ ( $p=0.000$ ). The pitch range of /se/ is significantly wider than that of /si/ in when /se/ is focused (Narrow-2).

However, the mean pitch range difference between /si/ and /se/ in Narrow-2 is significantly smaller than that in Narrow-1, supporting the observation that when the non-final /si/ is under narrow focus (Narrow-1), its pitch range is more expanded while that of the following /se/ is much more compressed. Such a substantial pitch range difference between /si/ and /se/ is reduced when the penultimate /se/ is focused (Narrow-2).

## 4. Discussions

The above findings illustrate that the pitch range of both target words is related to focus. This substantial pitch range expansion at focus, coupled with a significant increase in duration found in a pilot study [6], supports a similar finding reported by Eady & Cooper [7].

As discussed in Section 3.1, the sentence medial /si/ has a higher peak F0 than the penultimate /se/. Such lowering of F0 is particularly prominent when the High-Level (55) tone of /se/ is produced in citation, narrow focus (Narrow-2), and broad focus. Even when /se/ is focused, its peak F0 is not higher than that of the preceding non-focused /si/ (Figure 3). This phenomenon has been observed in tone languages like Mandarin [3]. One plausible explanation for this substantial drop in F0 topline comes from the fact that there are intervening low tones, i.e., Mid-Level tone (33), Mid-Low-Falling tone (21), and Mid-Low-Level (22) tone, between the two High-Level tones of /si/ and /se/. Under the framework proposed by Xu & Wang [8], a falling pitch movement is generally faster than a rising pitch movement. Given the same amount of time and effort, a rise cannot cover as much F0 range as a fall. Hence, a rise after a fall is unlikely to fully recover the drop in F0 resulting from the falling pitch movement unless an extra effort is given to the rise. As a result, the High-Level (55) tone of /si/ is higher than that of /se/. But when an extra effort is given to contrastively stress the target word /se/ (Narrow-2), such a substantial difference in peak F0 is reduced (Figure 3). This substantial drop in the F0 topline may also couple with final lowering observed in the study of Cantonese intonation by Vance [9].

Although a comparison of Figures 2 and 4 shows that the F0 pattern for narrow focus on /si/ resembles that for broad focus, a two-tailed *t* test does not show any significant pitch range difference between /si/ and /se/ in broad focus ( $p=0.08$ ). There is also no significant pitch range variation between the respective narrow focus and broad focus readings of /si/ and of /se/. One plausible explanation is the relatively heavy semantic weight carried by /si/ and /se/ in the carrier frame [2]. Therefore, when speakers were prompted to produce the test sentence in broad focus, *both* target words received “equal prominence” [10] in the frame.

If broad focus is considered to be the placement of focus on the whole intonation group [11] rather than on a whole sentence [2], the phenomenon of “equal prominence” can then be explained. The target words are located in the rightmost position of their respective intonation groups. As prominence is generally placed on the rightmost accentable item of an intonation group, the two target words will naturally be equally focused when the speakers were told to produce the test sentence in broad focus.

## 5. Conclusion

The findings of this study have revealed that other than final F0 lowering, focus and lexical tone play a crucial role in the formation of F0 contours in Cantonese declarative sentences. In particular, a substantial pitch range expansion is associated with focus regardless of sentence position, while a severe pitch range reduction is observed only after the focused item.

It has been suggested that speakers control the prosody of an utterance to signal linguistic information such as focus at the word or sentence level, and that listeners often parse the incoming speech signal first into prosodic rather than syntactic constituents [12]. The acoustic cues to prosodic phrase boundaries in Cantonese are yet to be determined, and due to a lack of unified prosodic transcription method, prosodic comparison across languages is often impossible. The prosodic transcription of data adopted in this study has thus put forward the importance of phonetically identifying prosodic phrases and boundaries in sentence prosody research.

Recent speech processing research has attempted to develop a method for tone recognition of isolated Cantonese syllables [13]. Since words are not uttered in isolation in speech, a tone recognition model for Cantonese connected speech is called for. The present F0 findings could provide a linguistic basis for building up such a model.

## 6. References

- [1] Halliday, M.A.K., 1970. *A course in spoken English: Intonation*. Oxford: Oxford University Press.
- [2] Ladd, R., 1980. *Intonational phonology*. Cambridge: Cambridge University Press.
- [3] Jin, S., 1996. *An acoustic study of sentence stress in Mandarin Chinese*. Ph.D. dissertation, The Ohio State University.
- [4] Xu, Y., 1999. Effects of tone and focus on the formation and alignment of F0 contours. *Journal of Phonetics*, 27, 55-105.
- [5] Cooper, W.; Eady, S.; Mueller, P., 1985. Acoustical aspects of contrastive stress in question-answer contexts. *Journal of the Acoustical Society of America*, 77(6), 2142-2155.
- [6] Man, Vicky C.H., 1999. *An acoustic study of the effects of sentential focus on Cantonese tones*. Unpublished M.A. thesis, University of Victoria.
- [7] Eady, S; Cooper, W., 1986. Speech intonation and focus location in matched statements and questions. *Journal of the Acoustical Society of America*, 80(2), 402-415.
- [8] Xu, Y.; Wang, E., 2001. Pitch targets and their realizations: Evidence from Mandarin Chinese. *Speech Communication*, 33, 319-337.
- [9] Vance, T., 1976. An experimental investigation of tone and intonation in Cantonese. *Phonetica*, 33, 368-392.
- [10] Rump, H.; Collier, R., 1996. Focus conditions and the prominence of pitch-accented syllables. *Language and Speech*, 39(1), 1-17.
- [11] Cruttenden, A., 1997. *Intonation*. Cambridge: Cambridge University Press.
- [12] Potisuk S., 1995. *Prosodic disambiguation in automatic speech understanding of Thai*. Ph.D. dissertation, Purdue University.
- [13] Lee, T.; Ching, P.C.; Chan, L.W.; Cheng, Y.H.; Mak, B., 1995. Tone recognition of isolated Cantonese syllables. In *Proceedings of the IEEE Transactions on Speech and Audio Processing*, 3(3), 204-209.