*Language Variation and Change*, **15** (2003), 211–225. Printed in the U.S.A.

DOI: 10.1017/S0954394503152039

# Variation and merger of the rising tones in Hong Kong Cantonese

ROBERT S. BAUER, CHEUNG KWAN-HIN, AND CHEUNG PAK-MAN
*Hong Kong Polytechnic University*

ABSTRACT

Two male speakers of Hong Kong Cantonese varied the endpoints of High Rising and Mid-Low Rising tones and merged them in both directions under experimental conditions. The variation and merger of the two rising tones raise the possibility that at least four tonal subsystems may coexist within the Hong Kong Cantonese speech community. Sociolinguistic research over the past 20 years has documented variation and change among Cantonese sound segments but not the tones. Tonal variation in Hong Kong Cantonese appears to be a potentially important sociolinguistic variable.

Among the world's tone languages Hong Kong Cantonese stands out as one with a particularly rich system of lexical tone contrasts. There are six relatively long contour tones on the live syllables (i.e., open syllables and syllables closed by nasals *-m*, *-n*, *-ŋ*) and three much shorter tones on the dead syllables (i.e., closed by the unreleased homorganic stops *-p*, *-t*, *-k*).[1] These tones have been numbered and described with Chao tone letters as follows (Bauer & Benedict, 1997:125):

| | | |
|---|---|---|
| T1 | High Level | ˥55 |
| T2 | Mid-Low Falling | ˩21 |
| T3 | High Rising | ˧˥25 |
| T4 | Mid-Low Rising | ˩˧23 |
| T5 | Mid Level | ˧33 |
| T6 | Mid-Low Level | ˩22 |
| T7 | High Stopped | ˥5 |
| T8 | Mid Stopped | ˧33 |
| T9 | Mid-Low Stopped | ˩22 |

As the Chao tone letters suggest, this is a relatively symmetrical system of level, falling, and rising tone contours, all neatly distributed within the tone space.[2]

This article is a revised version of a paper presented at the 33rd International Conference on Sino-Tibetan Languages and Linguistics in Trang, Thailand, on October 5, 2000. The research reported here was supported by the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. PolyU 5249/99H Linguistics) and by Hong Kong Polytechnic University Research Grant G-YB57.

Acoustic phoneticians generally agree that fundamental frequency (F0) is the major acoustic correlate of lexical tone. In an acoustic study of Hong Kong Cantonese tones, Kei et al. (2002) measured the tones produced by a sample of 56 male and female subjects. To compare tones across speakers they reduced inter-speaker variation due to physiological differences by normalizing the F0 values. They concluded that the acoustic results generally supported these verbal descriptions and symbolic representations of the Cantonese tones. One seemingly minor but potentially significant finding from this study was that six subjects merged the two rising tones, either to High Rising or to Mid-Low Rising. This type of change in the tones has not been previously reported for the standard variety of Hong Kong Cantonese and is rather surprising, given that these two rising tones are phonemically contrastive. It is quite understandable that Kei et al. (2002:29) described this merger as "tone production errors," as the purpose of their study was to produce a normative profile of the standard Cantonese tones for use by speech therapists. However, these "errors" can also be viewed as variant tones, which may differ from the standard tones in some systematic way that has yet to be identified.

The findings from Kei et al.'s acoustic study along with those from our own (Bauer et al., 2000) and anecdotal reports from colleagues all indicate that some younger Hong Kong Cantonese speakers are using only one rising tone. Variation among sound segments has been well documented for Hong Kong Cantonese in a number of sociolinguistic studies, but so far lexical tone has not been investigated as a sociolingustic variable. As Chambers (1997:197) noted, "whenever segmental and prosodic elements are involved in change, prosodic elements are likely to lag behind segmental ones. The reason for this is that prosodic features have primacy in the acquisition schedule … they are reinforced constantly … [and] are acquired very early."

We would like to suggest that Cantonese tonal variation deserves a place alongside variation in sound segments as a potentially interesting and important variable for sociolinguistic study. To shed more light on this matter we report our findings from two experiments on the variation and merger of Hong Kong Cantonese High Rising and Mid-Low Rising tones.

## EXPERIMENT 1: PRODUCTION OF CANTONESE TONES

This experiment—a small-scale replication of the acoustic study by Kei et al. (2002)—had as its purpose the measurement and analysis of the contrastive tones of Hong Kong Cantonese.

### *Subjects*

Participating in this experiment were eight male subjects who were all born and raised in Hong Kong; their ages ranged from 20 to 35 years with a mean age of 27 years. The social characteristics of the subjects were as follows: (1) LKC: 20 years old, second-year B.A. student; (2) WKW: 20 years old, second-year B.A.

student; (3) NCK: 20 years old, second-year B.A. student; (4) LWH: 29 years old, B.A. degree, university research assistant; (5) MHL: 29 years old, M.A. degree, professional translator; (6) WKK: 32 years old, B.A. degree, salesman; (7) TSW: 30 years old, Ph.D. degree, university assistant professor; (8) CCL: 35 years old, M.A. degree, university language instructor.

*Elicitation materials and recording procedures*

To elicit the subjects' production of the full set of contrastive Cantonese tones, we presented a list of 122 single Chinese characters (as opposed to combinations of two or more Chinese characters) to one group of six subjects, and a slightly shorter list to another group of two subjects. Each subject was instructed to read aloud the Chinese characters one by one. The subject sat before an AKG C451E microphone in a sound-treated room; his reading citation forms of these Chinese characters were recorded with a Panasonic SV 3700 DAT recorder at 44.1 kHz samples. Randomly distributed through the character lists were two minimal sets of Chinese characters whose pronunciations would elicit the six long Cantonese tones on the two open syllables jiː and jɐw (see Appendix, List 1, for the 12 Chinese characters that are associated with these two syllables; these were used in the elicitation). The first minimal set of six Chinese characters was read with open syllable jiː and the corresponding tones from an inventory of six phonemically contrastive tones. The second minimal set of six Chinese characters was read with open syllable jɐw and the corresponding tones. Each subject was recorded as he read through the list of Chinese characters five times to produce five pronunciations (or tokens) for each Chinese character.

*Analysis*

To facilitate acoustic analysis, the recordings were digitized at a sampling rate of 10 kHz by means of a 16-bit analog-to-digital (A/D) converter using the Kay Elemetrics Computerized Speech Lab (CSL) 4300B system. For the analysis of the subjects' productions of these tones, the second, third, and fourth tokens (labeled A, B, and C in the tables) were selected. The F0 was measured at six points along the F0 trajectory (0%, 20%, 40%, 60%, 80%, 100%) as displayed on the computer screen by the CSL. A tone contour's endpoint was the F0 value as measured at the 100% point (i.e., the end of the F0 trajectory).

*Results*

As was done in Kei et al., the *z*-score normalized F0 values were derived from F0, which was sampled at six points along the F0 trajectory, minus the set of mean F0 values and divided by the standard deviation based on the mean of the sampled points (see Rose, 1987).[3] The production of tones by the two subjects, TSW and CCL, seemed unusual in relation to the tones produced by the other six subjects, and so the tone data from TSW and CCL were not included in these calculations of tone normalization. The mean normalized F0 values for the six male subjects are listed in Table 1. These normalized F0 values become more meaningful when

TABLE 1. *Mean normalized F0 values (and standard deviations) of six tones for six male speakers*

| | Normalized Fundamental Frequency (F0) | | | | | |
|---|---|---|---|---|---|---|
| Tone | 0% | 20% | 40% | 60% | 80% | 100% |
| T1 | 1.52 (0.22) | 1.64 (0.26) | 1.68 (0.24) | 1.63 (0.25) | 1.59 (0.27) | 1.46 (0.30) |
| T2 | −0.18 (0.13) | −0.46 (0.17) | −0.92 (0.16) | −1.30 (0.24) | −1.65 (0.22) | −1.78 (0.52) |
| T3 | −0.29 (0.21) | −0.63 (0.32) | −0.80 (0.13) | −0.42 (0.31) | 0.67 (0.59) | 2.01 (0.40) |
| T4 | −0.36 (0.17) | −0.60 (0.24) | −0.79 (0.19) | −0.62 (0.20) | −0.17 (0.26) | 0.50 (0.09) |
| T5 | 0.34 (0.19) | 0.29 (0.25) | 0.15 (0.26) | 0.05 (0.21) | 0.01 (0.18) | 0.15 (0.14) |
| T6 | −0.06 (0.13) | −0.24 (0.12) | −0.44 (0.12) | −0.58 (0.16) | −0.71 (0.10) | −0.58 (0.34) |

*Note:* Percent values indicate duration points at which F0 was sampled.

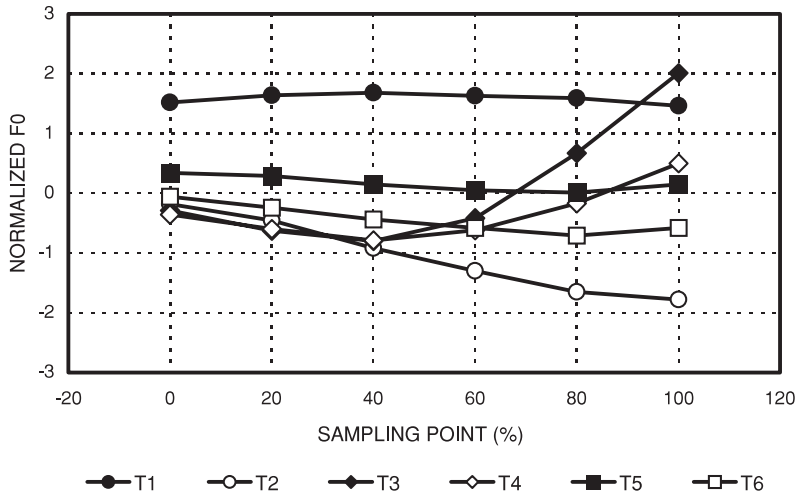



FIGURE 1. Mean *z*-score normalized F0 contours for six male speakers' tones. (T1 = High Level, T2 = Mid-Low Falling, T3 = High Rising, T4 = Mid-Low Rising, T5 = Mid Level, T6 = Mid-Low Level).

displayed as tone contours, as in Figure 1. From Figure 1 we observe that the relative shapes of most of the tone contours generally conform to their verbal descriptions. However, T1 High Level falls slightly at the end; both T5 Mid Level and T6 Mid-Low Level fall slightly in the middle and then rise slightly at the end. T3 High Rising terminates at a point somewhat higher than that of T1 High Level, even though they are both traditionally given the same High endpoint of 5. Also noteworthy is that the lower part of the figure is relatively crowded with four tones all sharing the same Mid-Low starting point of 2: namely, T2 ˩21, T3 ˧˥25, T4 ˧˥23, and T6 ˨˩22. There is no question that the standard variety of Hong

TABLE 2. *Normalized F0 values of highest points for High Level, High Rising, and Mid-Low Rising tones for eight male speakers*

| Subject | High Level | High Rising | Mid-Low Rising |
|---|---|---|---|
| LKC | 1.90 | 1.90 | 0.55 |
| WKW | 1.88 | 1.65 | 0.50 |
| NCK | 1.89 | 1.57 | 0.54 |
| LWH | 1.63 | 2.66 | 0.34 |
| MHL | 1.77 | 2.17 | 0.47 |
| WKK | 1.76 | 2.13 | 0.58 |
| Mean | 1.81 | 2.01 | 0.50 |
| TSW | 1.50 | 2.63 | 0.91 |
| CCL | 1.42 | 3.17 | 2.46 |

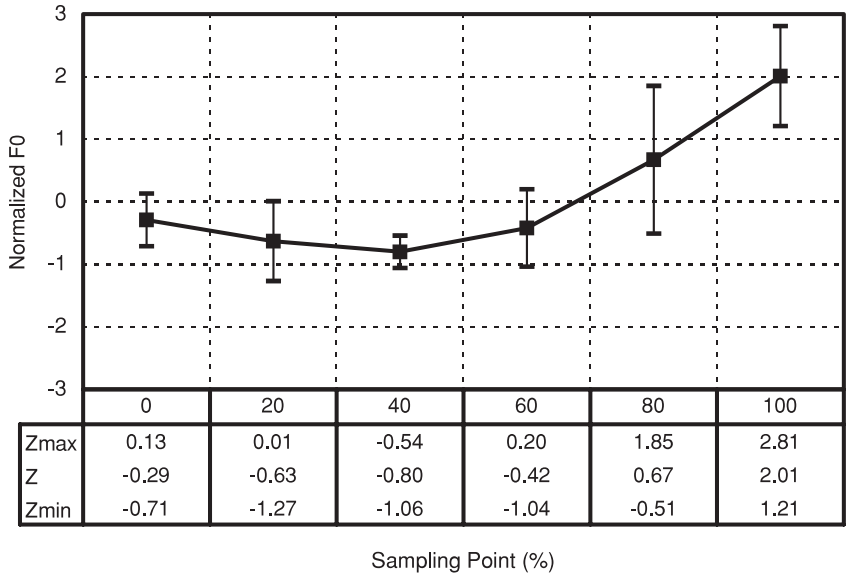
*Note:* Means based on six subjects.

Kong Cantonese phonemically contrasts the two rising tones. The important feature distinguishing these two tones is the substantial gap that separates their endpoints.[4]

In measuring the tone data produced by TSW and CCL for this experiment, we observed that the F0 peaks (endpoints) of their rising tones seemed to be anomalous in relation to the F0 peaks of their High Level tones as well as the peaks of the High Level, High Rising, and Mid-Low Rising tones of the other six subjects. To compare the endpoints of the two rising tones for all eight subjects and to ascertain a "normal" range, we normalized the F0 values for the six sampling points and then compared the normalized peaks of the High Level, High Rising, and Mid-Low Rising tones. Table 2 lists the normalized F0 values of the peaks of these tones for all eight speakers. In Table 2 the subjects are divided into two groups according to their production of the two rising tones: the first group (six subjects) represented the tones of the standard variety, and the second group (TSW and CCL) has tone productions that deviate from the norm.

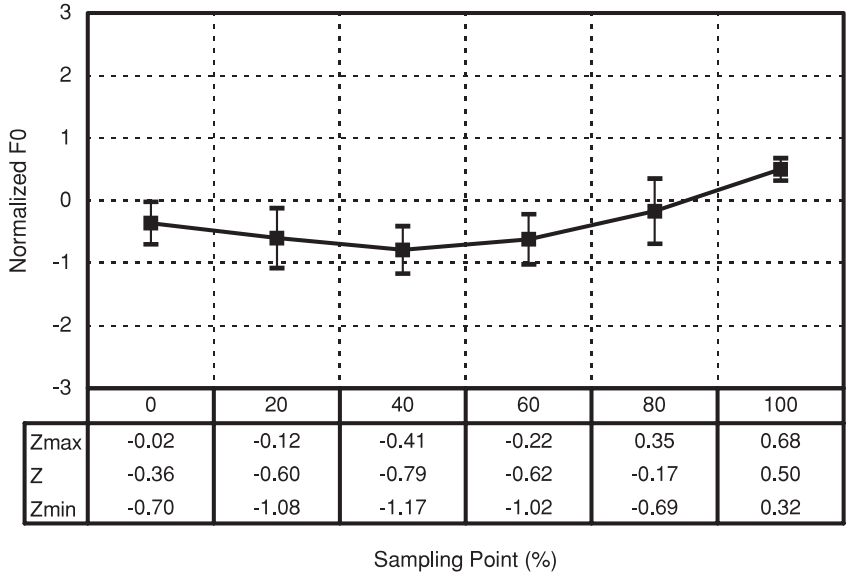
Figures 2 and 3 plot the mean normalized F0 values for the High Rising and Mid-Low Rising tones for the first group, along with two standard deviations above and below the mean ($Z_{max}$, $Z_{min}$). A range of four standard deviations around the mean encompasses roughly 95% of all normally distributed observations; this is the range of variation we would expect for the normalized F0 values of the six lexical tones in Cantonese.

On the basis of the combined information in Table 2 and Figures 2 and 3, we can compare the *z*-score normalized F0 values of TSW and CCL with the mean and $Z_{max}$ scores for the six subjects whose tones were considered to be normal. The endpoint of TSW's Mid-Low Rising tone had a mean *z* score of 0.91, which was higher than the mean of 0.50 and the $Z_{max}$ of 0.68; this difference suggests that his production of the tone was anomalous. TSW's High Rising tone had a mean *z* score of 2.63, which was above the mean of 2.01 but still below the $Z_{max}$




| | 0 | 20 | 40 | 60 | 80 | 100 |
|---|---|---|---|---|---|---|
| Zmax | 0.13 | 0.01 | -0.54 | 0.20 | 1.85 | 2.81 |
| Z | -0.29 | -0.63 | -0.80 | -0.42 | 0.67 | 2.01 |
| Zmin | -0.71 | -1.27 | -1.06 | -1.04 | -0.51 | 1.21 |

FIGURE 2. Mean *z*-score normalized F0 values for High Rising tone (T3) based on data for six subjects. Vertical bars indicate two standard deviations above and below mean.




| | 0 | 20 | 40 | 60 | 80 | 100 |
|---|---|---|---|---|---|---|
| Zmax | -0.02 | -0.12 | -0.41 | -0.22 | 0.35 | 0.68 |
| Z | -0.36 | -0.60 | -0.79 | -0.62 | -0.17 | 0.50 |
| Zmin | -0.70 | -1.08 | -1.17 | -1.02 | -0.69 | 0.32 |

FIGURE 3. Mean *z*-score normalized F0 values for Mid-Low Rising tone (T4) based on data for six subjects. Vertical bars indicate two standard deviations above and below mean.

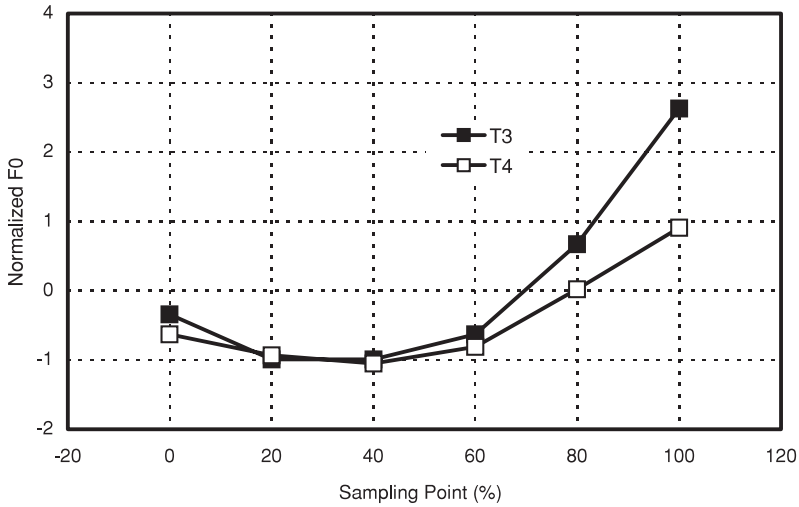



FIGURE 4. Mean *z*-score normalized F0 values for TSW's High Rising tone (T3) and Mid-Low Rising tone (T4).

of 2.81, and so it was considered to be normal. The endpoints of CCL's High Rising and Mid-Low Rising tones (3.17 and 2.46) were very much higher than the means of 2.01 and 0.50 and the $Z_{max}$s of 2.81 and 0.68; the numbers once again indicate that CCL's production of the two rising tones deviated from the norm. To show the anomaly of the two speakers' rising tones, Figures 4 and 5 display the mean normalized F0 values for the High Rising and Mid-Low Rising tones of TSW and CCL, respectively.

*Discussion*

Upon closer examination of the acoustic data, the anomaly of TSW's Mid-Low Rising tone, with its higher than normal endpoint, was found to be associated with his pronunciations of ji˧˩23 'ear'. Table 3 lists the F0 endpoints of his High Rising and Mid-Low Rising tones for the two sets of minimal pairs from the first experiment. (The pronunciation of the Chinese characters in the standard variety of Hong Kong Cantonese is represented here with phonemic IPA transcription.)

In TSW's pronunciation of 'ear' the Mid-Low Rising tone had endpoints of 127 Hz, 139 Hz, and 143 Hz for a mean of 136.33 Hz. The mean F0 endpoint of his High Rising tone for three tokens of 'chair' and three tokens of 'pomelo' was 152.50 Hz. The mean endpoint of the Mid-Low Rising tone, based on three tokens of 'have', was 112.33 Hz. That is, he had a 40 Hz difference between the mean endpoints of his High Rising and Mid-Low Rising tones. But the endpoints of 'ear', with a mean value of 136.33 Hz, was somewhere in between these two values, so that they resembled neither the High Rising nor the Mid-Low Rising

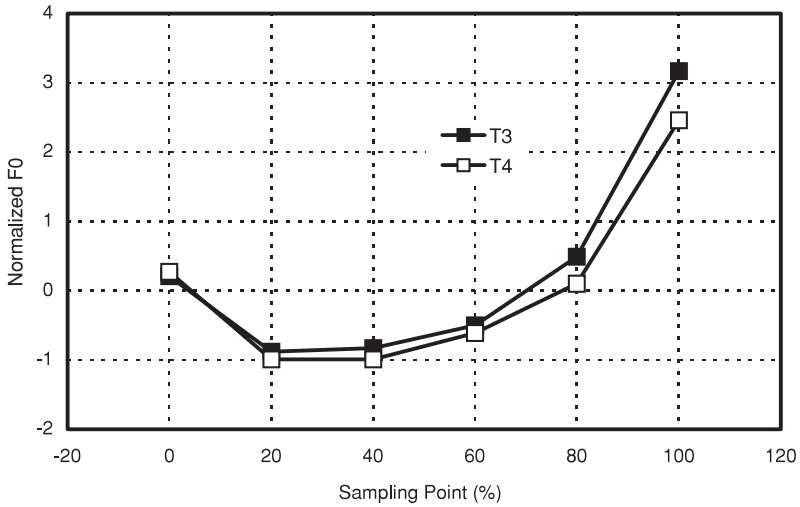



FIGURE 5. Mean *z*-score normalized F0 values for CCL's High Rising tone (T3) and Mid-Low Rising tone (T4).

TABLE 3. *Comparison of mean endpoints of High Rising and Mid-Low Rising tones for two minimal pairs by TSW*

| High Rising | | | Mid-Low Rising | | |
|---|---|---|---|---|---|
| Word | Token | Tone Endpoint (Hz) | Word | Token | Tone Endpoint (Hz) |
| 'chair' jiː˨˦25 | A | 147 | 'ear' jiː˨˩23 | A | 127 |
| | B | 152 | | B | 139 |
| | C | 143 | | C | 143 |
| Mean | | 147.33 | | | 136.33 |
| 'pomelo' jɐw˨˦25 | A | 152 | 'have' jɐw˨˩23 | A | 111 |
| | B | 169 | | B | 112 |
| | C | 152 | | C | 114 |
| Mean | | 157.67 | | | 112.33 |
| Overall mean | | 152.50 | | | 124.33 |

tones. This suggests that he was unsure about which rising tone was the appropriate one for 'ear'.

As for CCL, his F0 endpoints for the two rising tones were both unusual in that they considerably exceeded the group mean and $Z_{max}$ scores. His High Rising tone endpoint was 3.17, which was very high (in contrast, the peak of his High Level tone was only 1.42, well below the group mean of 1.81); however, his Mid-Low Rising tone endpoint was 2.46, which actually resembled the High Rising tone of some of the six subjects. Figure 5 displays CCL's two rising tones

on the same graph; their endpoints came so close together that we had to conclude that he was not contrastively producing them. Reading aloud the Chinese characters may have put CCL in the position of having to confront for the first time his use of the rising tone. It would appear from his tone production that he only used the High Rising tone; perhaps he realized from the structure of the experiment that he was supposed to be using two different tones but felt uncertain about what these were and how to produce them.

At any rate, we decided to investigate further the production of the rising tones by TSW and CCL to determine whether they were able to differentiate or contrast the two tones for other lexical items. To do this we devised a second experiment with a set of minimal pairs for which the High Rising and Mid-Low Rising tones contrast in the standard variety of Cantonese.

## EXPERIMENT 2: PRODUCTION OF RISING TONES BY TWO SUBJECTS

There were two questions about the production of the two rising tones by TSW and CCL that we wanted to answer. First, was TSW's indecision about the choice of rising tones limited to ji: ˩˧23 'ear' or would it occur with other items belonging to either the High Rising or the Mid-Low Rising tone categories? Second, was CCL able to contrast the two rising tones?

### *Elicitation materials*

These two subjects were presented with a list of six minimal pairs of Chinese characters that contrast the two rising tones. In this experiment each subject was given a sheet of paper on which the pairs appeared in two columns (see Appendix, List 2). The tonal contrasts were in a random order: that is, the character on the left took the High Rising tone and the one on the right took the Mid-Low Rising tone or vice versa. This experiment required the subject to read each character aloud and to decide if the pair had the same or different tone.

### *Results and discussion*

Table 4 lists the endpoints of TSW's High Rising and Mid-Low Rising tones for the 36 tokens (3 tokens × 12 lexical items). The asterisks in the far right column indicate that the F0 difference between the endpoints of High Rising and Mid-Low Rising tones is zero or negative. As this table shows, the mean endpoints for the two sets of lexical items are 103.39 Hz for the High Rising tone and 104.28 Hz for the Mid-Low Rising tone: that is, these two tones have essentially identical endpoints. At the same time, these figures are somewhat lower than that for 'have' (112.33 Hz) in the first experiment. On the basis of these findings, we had to conclude that, in the second experiment, TSW produced only Mid-Low Rising tone contours for the 18 tokens of the items in the High Rising tone category as well as for the 18 tokens of the items in the Mid-Low Rising tone category.

In the first experiment, TSW did distinguish between High Rising and Mid-Low Rising tones by producing tone contours with obviously different endpoints,

TABLE 4. *Mean endpoints of High Rising and Mid-Low Rising tones (in Hz) for six minimal pairs of words, as read by TSW*

| High Rising | | | Mid-Low Rising | | | |
|---|---|---|---|---|---|---|
| Word | Token | Tone Endpoint (Hz) | Word | Token | Tone Endpoint (Hz) | F0 Difference |
| 'history' siː˨˧25 | A | 119 | 'market' siː˨23 | A | 114 | 5 |
| | B | 98 | | B | 108 | −10* |
| | C | 87 | | C | 96 | −9* |
| Mean | | 101 | | | 106 | −5* |
| 'write' sɛː˨˧25 | A | 111 | 'society' sɛː˨23 | A | 114 | −3* |
| | B | 90 | | B | 107 | −17* |
| | C | 97 | | C | 101 | −4* |
| Mean | | 99 | | | 107 | −8* |
| 'powder' fɐn˨˧25 | A | 107 | 'struggle' fɐn˨23 | A | 109 | −2* |
| | B | 105 | | B | 100 | 5 |
| | C | 97 | | C | 103 | −6* |
| Mean | | 103 | | | 107 | −4* |
| 'tolerate' jɐn˨˧25 | A | 99 | 'guide' jɐn˨23 | A | 96 | 3* |
| | B | 92 | | B | 101 | −9* |
| | C | 100 | | C | 97 | 3* |
| Mean | | 97 | | | 98 | −1* |
| 'smallpox' maː˨˧25 | A | 121 | 'horse' maː˨23 | A | 109 | 12 |
| | B | 109 | | B | 102 | 7 |
| | C | 110 | | C | 101 | 9 |
| Mean | | 113 | | | 104 | 9 |
| 'tiger' fuː˨˧25 | A | 119 | 'married woman' fuː˨23 | A | 115 | 4 |
| | B | 103 | | B | 103 | 0* |
| | C | 97 | | C | 101 | −4* |
| Mean | | 106 | | | 106 | 0* |
| Overall mean | | 103.39 | | | 104.28 | −0.89* |

*Note:* * marks 0 or negative F0 value difference.

but in the second experiment, which required him to contrast the two tones, he did not differentiate them but switched over to the Mid-Low Rising tone contour. It would appear that TSW had decided that the correct tone to use for all the lexical items in the second experiment was the Mid-Low Rising tone, and so he merged (neutralized) his High Rising tone to Mid-Low Rising.

As described for CCL, the mean endpoint of his Mid-Low Rising tone was about 13 Hz less than that of his High Rising tone. Nonetheless, a comparison of the actual endpoint values for both sets of tokens shows that they were very similar for the two tones. With respect to the four Chinese characters we used in the first elicitation experiment, we initially conjectured that CCL tried to contrast the High Rising and Mid-Low Rising tones, but was unable to distinguish between them in a stable and consistent manner and therefore produced the two tone contours more or less randomly.

To pursue this hypothesis further, CCL was presented with the same set of minimal pairs as TSW. In this experiment, CCL was faced with the task of distinguishing the Chinese characters solely on the basis of their tones. In Table 5 we list the F0 endpoints of his High Rising and Mid-Low Rising tones. As in Table 4,

TABLE 5. *Mean endpoints of High Rising and Mid-Low Rising tones (in Hz) for six minimal pairs of words, as read by CCL*

| High Rising | | | Mid-Low Rising | | | |
|---|---|---|---|---|---|---|
| Word | Token | Tone Endpoint (Hz) | Word | Token | Tone Endpoint (Hz) | F0 Difference |
| 'history' siː˨˥25 | A | 154 | 'market' siː˨˧23 | A | 154 | 0* |
| | B | 154 | | B | 152 | 2 |
| | C | 192 | | C | 161 | 31 |
| Mean | | 167 | | | 156 | 11 |
| 'write' sɛː˨˥25 | A | 128 | 'society' sɛː˨˧23 | A | 185 | −57* |
| | B | 161 | | B | 167 | −6* |
| | C | 164 | | C | 173 | −9* |
| Mean | | 151 | | | 175 | −24* |
| 'powder' fɐn˨˥25 | A | 125 | 'struggle' fɐn˨˧23 | A | 152 | −27* |
| | B | 200 | | B | 167 | 33 |
| | C | 147 | | C | 179 | −32* |
| Mean | | 157 | | | 166 | −9* |
| 'tolerate' jɐn˨˥25 | A | 139 | 'guide' jɐn˨˧23 | A | 149 | −10* |
| | B | 135 | | B | 152 | −17* |
| | C | 154 | | C | 189 | −35* |
| Mean | | 143 | | | 163 | −20* |
| 'smallpox' maː˨˥25 | A | 122 | 'horse' maː˨˧23 | A | 166 | −44* |
| | B | 166 | | B | 170 | −4* |
| | C | 182 | | C | 169 | 13 |
| Mean | | 157 | | | 168 | −11* |
| 'tiger' fuː˨˥25 | A | 159 | 'married woman' fuː˨˧23 | A | 139 | 20 |
| | B | 164 | | B | 154 | 10 |
| | C | 213 | | C | 147 | 66 |
| Mean | | 179 | | | 147 | 32 |
| Overall mean | | 158.83 | | | 162.50 | −4.17 |

*Note:* * marks 0 or negative F0 value difference.

the asterisks indicate that the F0 difference between the endpoints of High Rising and Mid-Low Rising tones is zero or negative.

As this tables shows, for CCL's High Rising tone the endpoints range from 122 Hz to 213 Hz with a mean of 158.83 Hz; for his Mid-Low Rising tone the endpoints range from 139 Hz to 189 Hz with a mean of 162.50. The mean value of the endpoint for his Mid-Low Rising tone (162.50 Hz) is slightly higher (4.17 Hz) than that for his High Rising tone (158.83 Hz). On the basis of the results from the two experiments, we concluded that CCL did not have a Mid-Low Rising tone. He typically produced the High Rising tone when pronouncing Chinese characters that belong to both the High Rising and Mid-Low Rising tone categories. Two alternative questions would seem to arise here. Was there a time when CCL had two different, contrastive rising tones and then later merged them into High Rising? Or, had he only had the High Rising tone all along? Either situation seems plausible. However, in view of tonal developments in other Cantonese dialects of the neighboring region, we would tentatively answer "yes" to the second question.

## CONCLUSION

As part of their acoustic study of the Hong Kong Cantonese tones, Kei et al. (2002:28–29) conducted a perceptual analysis of the tones produced by a group of 15 adult subjects (9 males and 6 females) who had a mean age of 33.5 years. The experimental group included seven undergraduate students from the University of Hong Kong, and eight professionals who were either audiologists or teachers of the deaf. Kei et al. discovered that six subjects did not distinguish the High Rising and Mid-Low Rising tones in their production of tones according to standard Cantonese. Of these, two subjects realized the High Rising tone as Mid-Low Rising, three subjects realized Mid-Low Rising as High Rising, and one subject produced tone contours with endpoints that were midway between High Rising and Mid-Low Rising. They described their subjects' deviations from the norm as "tone production errors" (Kei et al., 2002:29). However, another way of looking at the tones produced by these subjects would be to say that two subjects merged the rising tones to Mid-Low Rising and three subjects merged them to High Rising. The sixth subject produced one rising tone with a contour that terminated halfway between Mid-Low Rising and High Rising. Their description of these tone mergers mirrors very closely our own findings of the tone productions of CCL and TSW.

As stated in So (1996:188), So and Varley (1991) found from their own tone perception study that some subjects in a group of 101 Hong Kong Cantonese speakers "often confused the High [Rising] and [Mid-Low Rising] tones." The reason given for this confusion was that the two tones have "similar starting points" (as summarized in So, 1996:188). But the fact that the two tones begin at the same point would hardly seem to be a likely explanation in view of the findings from the tone perception study by Kei et al., which found that judges relied on the difference between the endpoints of the two tone contours as the salient cue for perceptually identifying them.

Just as we cannot assume that every speaker of the speech community has exactly the same inventory of contrastive sound segments, the same must be said about the set of contrastive tones that make up the tone system. Not long after we completed this study, we recorded a few other young Hong Kong-born speakers. We found that they did have two rising tones but seemed to use them interchangeably as nondistinctive variants. Clearly, some kind of change is going on in the Hong Kong Cantonese tone system.

Findings from the present study as well as those from Kei et al. (2002) clearly reveal that some Hong Kong Cantonese speakers do not distinguish between the two rising tones, but tend to favor one or the other or even one in between the two. According to Hombert (1978:104), "[i]f two tones are too similar phonetically, they can either move away from each other in the tone space or merge, as happened, for instance, in Vietnamese, Lahu, and numerous Chinese dialects." Although he did not cite any specific Chinese dialects in which such a merger has occurred, Hong Kong Cantonese may be one example.

Within the Hong Kong Cantonese speech community we may be able to distinguish at least three subvarieties of speech based on tonal subsystems in which

the difference between these two rising tones is not maintained as in the standard variety. In some subvarieties the two tones have merged to High Rising, in others they have merged to Mid-Low Rising, and in still others the height of the rising tone varies between these two extremes. The notion of subvarieties based on differences within closely related tone systems can be viewed within the perspective of diachronic change and synchronic variation. During the historical development of standard Cantonese, the representative variety of the Yue group of dialects, the second historical tone category known as *Shang* 'rising', split into two other tone categories; in the traditional terminology these are called *Yin Shang* 'upper rising' and *Yang Shang* 'lower rising' and correspond to the terms High Rising and Mid-Low Rising, respectively. However, in several Yue dialects of the neighboring Pearl River Delta region in Guangdong province, the historical *Shang* tone category did not split, so there is only one tone contour associated with this tone category.[5]

Tonal variation in contemporary Hong Kong Cantonese has become the subject of comment as well as the object of linguistic study. If variation and change of the rising tones may be sociolinguistic in nature, how extensive is this tonal variation across the community? What kind of speakers have only the High Rising tone or only the Mid-Low Rising tone or both rising tones in phonemic contrast? Over the past 20 years various types of studies (descriptive, sociolinguistic, perceptual) have been documenting variation and change of sound segments in Hong Kong Cantonese: for example, the mergers of standard Cantonese nasal syllabic *ŋ* to *m* and labialized velar initial *kw-* to the plain velar *k-* before the back round vowel *-ɔː* (Bauer, 1982, 1983); the merger of syllable-final velar nasal and stop *-ŋ/-k* to alveolar nasal and stop *-n/-t*, respectively (Bauer, 1979; Cheung Pak-man, 1998; Law et al., 2001).

From the point of view of areal linguistics, none of these segmental changes in Hong Kong Cantonese are isolated phenomena: their corresponding counterparts have occurred in some neighboring Yue dialects as well.[6] Variation and merger of the Hong Kong Cantonese rising tones may turn out to be similar to the sociolinguistically-based variation and merger of Cantonese sound segments. The relationship between High Rising and Mid-Low Rising tones within the Hong Kong Cantonese speech community appears to be a potentially interesting and important sociolinguistic variable that merits more investigation.

## NOTES

**1.** The terms "live" and "dead" syllables, which we borrow from Tai linguistics, provide a convenient terminology to refer to the co-occurrence of syllables with certain types of endings with long and short tones (see Bauer & Benedict, 1997:115).

**2.** In addition to the six tones listed here, there is also the High Falling tone ˥˩52, which occurs as a free variant of the High Level tone in the speech of some Hong Kong Cantonese speakers, for a total of seven long phonetic tones. However, in the speech of educated Cantonese speakers in Guangzhou, the provincial capital of Guangdong, which is located about 150 km north of Hong Kong, the High Falling tone is the default tone and the High Level tone is a word-derivational device that marks familiar, concrete nouns (see Bauer & Benedict, 1997:165–234).

A comparison of the Chao tone letters presented here with some previous acoustic studies finds disagreement about the starting point of the High Rising tone. Both Yue-Hashimoto (1972:92) and Vance (1977:96) assigned Chao tone letter ˧˥35 to this tone. (Vance noted on p. 103 that his own data indicated the starting point was closer to 2.) The subsequent studies cited here have clearly established that the High Rising tone along with three other tones all have the same Mid-Low starting

point of 2. Rose (1996:309) accepted 2 as the starting point of High Rising, but claimed its endpoint was Mid High (or 4). However, his own F0 data, presented as displays of the Cantonese tone contours in Figures 21-1 and 21-2 (A) (pp. 309, 312, respectively), clearly show the endpoints of the High Rising tone contour coming very close to or actually meeting the endpoints of the High Level tone contour, so his choice of 4 instead of 5 to symbolize this tone's terminal height does not seem to be justified.

**3.** The formula for calculating F0 normalization is as follows: $F0_{norm} = (F0_i - x)/SD$, where $F0_i$ is a sampling point, $x$ is the mean F0 from all sampling points, and $SD$ is the standard deviation around the mean of those points.

**4.** Contrary to Killingley (1985:31), the Mid-Low Rising tone is not a phonetic variant of the High Rising tone in the standard variety of Hong Kong Cantonese (as it seems to be in her Malaysian Cantonese); it is a regular "phonological tone" (in her terminology). Now, as she pointed out, we may not be able to find numerous free-standing monosyllabic words (i.e., free morphemes of one syllable) on which all six of the long tones contrast, but we can assemble a number of minimal sets of six Chinese characters, all of which are pronounced with one syllable but are differentiated by the six distinctive tones, as shown in List 1 of the Appendix. Furthermore, the Mid-Low Rising tone occurs on a number of important, monosyllabic free morphemes (or words) with very high frequency in speech: for example, the pronouns ŋɔ̌ː 'I', nêj 'you', and khθ̂ɥ 'he, she' and the existential verbs jêw 'have' and môw 'not have', among others.

**5.** For example, the *Shang* tone contour is rising in Kamtin in the New Territories of Hong Kong, Macao, Conghua City, Gaoming, Zhongshan (Shiqi), and Zhuhai; but it is High Level ⌉55 in Enping and Low Level ⌋11 in Baoan according to Zhan and Cheung (1987:9, 10, 12, 17, 18, 23, 26).

**6.** The change of syllabic ŋ to m has also taken place in the dialects of Macao, Conghua, Zengceng, Doumen, Enping, and Dongguan, whereas the Xinhui and Taishan dialects are reported to have variation between the two nasal syllabics. In a few dialects of the region the velar nasal and stop endings have alveolarized (become -n and -t). Bauer and Benedict (1997:336–339) briefly summarized these sound changes that have been documented in the Yue dialects of the Pearl River Delta region by Zhan and Cheung (1987).

## REFERENCES


Bauer, Robert S. (1979). Alveolarization in Cantonese: a case of lexical diffusion. *Journal of Chinese Linguistics* 7.1:132–141.

_____ (1982). *Cantonese sociolinguistic patterns: Correlating social characteristics of speakers with phonological variables in Hong Kong Cantonese*. Doctoral dissertation, University of California, Berkeley. Ann Arbor: University Microfilms International.

_____ (1983). Cantonese sound change across subgroups of the Hong Kong speech community. *Journal of Chinese Linguistics* 11:303–356.

Bauer, Robert S., & Benedict, Paul K. (1997). *Modern Cantonese phonology*. Berlin: Mouton de Gruyter.

Bauer, Robert S., Cheung, Kwan-hin, & Cheung, Pak-man. (2000, December). *An acoustic description of lexical tones in Hong Kong Cantonese*. Paper presented at Annual Research Forum, Linguistic Society of Hong Kong, City University of Hong Kong.

Chambers, J. K. (1997). *Sociolinguistic theory*. Oxford: Blackwell.

Cheung, Kwan-hin. (1986). *Phonology of present-day Cantonese*. Doctoral dissertation, University College, London.

Cheung, Pak-man. (1998). Merging of consonantal endings in Hong Kong Cantonese: A microhistoric study. Unpublished manuscript.

Hombert, Jean-Marie. (1978). Consonant types, vowel quality, and tone. In Victoria A. Fromkin (ed.), *Tone: A linguistic survey*. New York: Academic. 77–111.

Kei, Joseph, Smyth, Veronica, So, Lydia K. H., Lau, C. C., & Capell, Ken. (2002). Assessing the accuracy of production of Cantonese lexical tones: A comparison between perceptual judgement and an instrumental measure. *Asia Pacific Journal of Speech, Language and Hearing* 7:25–38.

Killingley, Siew-Yue. (1985). *A new look at Cantonese tones, five or six?* New Castle upon Tyne: Grewatt & Grewatt.

Law, Sam-Po, Fung, Roxana S-Y., & Bauer, Robert S. (2001). Perception and production of Cantonese consonant endings. *Asia Pacific Journal of Speech, Language and Hearing* 6:179–195.

Rose, Philip (1987). Consideration in the normalization of the fundamental frequency of linguistic tone. *Speech Communication* 6:343–351.

_____ (1996). Between- and within-speaker variation in the fundamental frequency of Cantonese citation tones. In Pamela Davis & Neville H. Fletcher (eds.), *Vocal fold physiology: Controlling complexity and chaos*. San Diego: Singular Publishing Group. 307–324.

So, Lydia K. H. (1996). Tonal changes in Hong Kong Cantonese. *Current Issues in Language & Society* 3:186–189.

So, Lydia K. H., & Varley, R. (1991). *Cantonese Lexical Comprehension Test*. Hong Kong: Department of Speech and Hearing Sciences, University of Hong Kong.

Vance, Timothy. (1977). Tonal distinctions in Cantonese. *Phonetica* 34:93–107.

Yue-Hashimoto, Ann. (1972). *Studies in Yue dialects 1: Phonology of Cantonese*. Cambridge: Cambridge University Press.

Zhan, Bohui, & Cheung, Yat-Shing. (1987). *A survey of dialects in the Pearl River Delta*, Vol. 1, *Comparative morpheme-syllabary* [in Chinese]. Hong Kong: New Century Publishing House.

## APPENDIX

LIST 1. *Chinese characters used in the elicitation of the six long contrastive tones of Hong Kong Cantonese in Experiment 1*

| T1. High Level | T3. High Rising | T5. Mid Level |
|---|---|---|
| 衣 jiː˥55 ‘clothes’ | 椅 jiː˧˥25 ‘chair’ | 意 jiː˧33 ‘idea’ |
| 休 jɐw˥55 ‘rest’ | 柚 jɐw˧˥25 ‘pomelo’ | 幼 jɐw˧33 ‘thin’ |
| **T2. Mid-Low Falling** | **T4. Mid-Low Rising** | **T6. Mid-Low Level** |
| 疑 jiː˩˧21 ‘suspicious’ | 耳 jiː˧˩23 ‘ear’ | 二 jiː˨22 ‘two’ |
| 油 jɐw˩˧21 ‘oil’ | 有 jɐw˧˩23 ‘have’ | 右 jɐw˨22 ‘right (side)’ |

LIST 2. *Chinese characters used in the elicitation of High Rising and Mid-Low Rising tones in Experiment 2*

| T3. High Rising: | T4. Mid-Low Rising: |
|---|---|
| 史 siː˧˥25 ‘history’ | 市 siː˧˩23 ‘market’ |
| 寫 sɛː˧˥25 ‘write’ | 社 sɛː˧˩23 ‘society’ |
| 粉 fɐn˧˥25 ‘powder’ | 奮 fɐn˧˩23 ‘struggle’ |
| 忍 jɐn˧˥25 ‘tolerate’ | 引 jɐn˧˩23 ‘guide’ |
| 痲 maː˧˥25 ‘smallpox’ | 馬 maː˧˩23 ‘horse’ |
| 虎 fuː˧˥25 ‘tiger’ | 婦 fuː˧˩23 ‘married woman’ |