

# The Prosodic Characteristics of the Number Words in Hong Kong Cantonese

Wai-Sum Lee

Department of Linguistics  
University of Hong Kong, Hong Kong  
wsleebea@hku.hk

## Abstract

The paper investigates the prosodic characteristics, including the temporal structure,  $F_0$ , and intensity, of the monosyllabic number words (MNW) and compound number words (CNW), bisyllabic and trisyllabic, in Hong Kong Cantonese. Results of analysis show that (i) the rime durations of all the MNW in isolation are reduced when the MNW occur as component MNW in the CNW; (ii) not all the initial consonants of the component MNW in the CNW are reduced; (iii) the reduction of the rimes and initial consonants for the component MNW varies according to the position in the CNW; (iv) the closure duration in the CNW also varies according to the position in the CNW; (v) the shapes of the  $F_0$  contours of the citation tones on the component MNW in the CNW are maintained, though the  $F_0$  level is slightly lowered for the MNW in the CNW-final position; and (vi) the intensity level is also slightly lowered for the component MNW in the CNW-final position.

## 1. Introduction

The prosodic characteristics, such as duration,  $F_0$ , and intensity, of the rimes of any component syllables of a bisyllabic or polysyllabic compound word are expected to differ from those of the same syllables occurring in isolation as monosyllabic words. It is also expected that the same component syllables occurring in different positions in a compound word also differ from each other in prosody, which in fact has been reported in a number of studies ([1, 2, 4, 5]). This paper investigates the differences in duration,  $F_0$ , and intensity, between the component *monosyllabic number words* (henceforth, MNW) that occur in the different positions in the *compound number words* (henceforth, CNW), bisyllabic and trisyllabic, in Hong Kong Cantonese (henceforth, HKC).

The number words in HKC may be monosyllables or compounds. In the case of the CNW, they are formed by combining members of a set of MNW, e.g., ‘eleven’ [sep<sup>2</sup> jɛt<sup>5</sup>] = ‘ten’ [sep<sup>2</sup>] + ‘one’ [jɛt<sup>5</sup>], and ‘twenty-one’ [ji<sup>22</sup> sep<sup>2</sup> jɛt<sup>5</sup>] = ‘two’ [ji<sup>22</sup>] + ‘ten’ [sep<sup>2</sup>] + ‘one’ [jɛt<sup>5</sup>]. Thus, a comparison of the MNW and the CNW, bisyllabic and polysyllabic, will show (i) the differences in prosody between the rimes of the MNW and those of the CNW, and (ii) the prosodic characteristics of the number words of different lengths.

In this study, the number words from ‘one’ to ‘ninety-nine’ in HKC were analyzed for their prosodic characteristics. The 99 test number words are formed by combining the ten basic MNW which are ‘one’ [jɛt<sup>5</sup>], ‘two’ [ji<sup>22</sup>], ‘three’ [sam<sup>55</sup>], ‘four’ [sei<sup>33</sup>], ‘five’ [m<sup>23</sup>], ‘six’ [lok<sup>2</sup>], ‘seven’ [ts<sup>h</sup>et<sup>5</sup>], ‘eight’ [pat<sup>3</sup>], ‘nine’ [kəu<sup>2</sup>], and ‘ten’ [sep<sup>2</sup>]. Thus, the number words from ‘one’ to ‘ten’ are monosyllabic. The number words from ‘eleven’ to ‘nineteen’ are bisyllabic, containing the MNW ‘ten’ [sep<sup>2</sup>] to be followed by one of the other nine MNW. The number words ‘twenty’, ‘thirty’, ‘forty’, ‘fifty’, ‘sixty’, ‘seventy’, ‘eighty’, and ‘ninety’ are also bisyllabic CNW which contain the MNW ‘ten’ [sep<sup>2</sup>] as the final syllable preceded by one of the other nine MNW. All the remaining number words from ‘twenty-one’ to ‘ninety-nine’ are trisyllabic CNW, which contain the MNW ‘ten’ [sep<sup>2</sup>] as the second or medial component syllable preceded and followed by one of the other nine MNW.

## 2. Method

The randomized lists of the 10 monosyllabic, 17 bisyllabic, and 72 trisyllabic test number words were read by 4 native speakers of HKC, 2 male and 2 female. The speakers, born and grew up in Hong Kong, were all university students in their early twenties. The recordings were performed in a sound proof booth (IAC). The speech analysis software CSL4400 of Kay Elemetrics was used for the analysis. The durations of the initial consonant and the rime of each component syllable in the test number words were measured directly from the speech waveforms with reference to the synchronized wideband spectrograms of the speech signals. The closure period if any between any successive component syllables of the bisyllabic and trisyllabic CNW was also measured. The  $F_0$  and intensity contours of the component syllables of the test number words were obtained, using the pitch synchronous method provided by the software.

## 3. Results

### 3.1. Temporal structure

#### 3.1.1. Rime duration

Tables 1a-1d show the mean durations of the rimes of the 10 MNW (‘one’ [jɛt<sup>5</sup>], ‘two’ [ji<sup>22</sup>], ‘three’ [sam<sup>55</sup>], ‘four’ [sei<sup>33</sup>], ‘five’ [m<sup>23</sup>], ‘six’ [lok<sup>2</sup>], ‘seven’ [ts<sup>h</sup>et<sup>5</sup>], ‘eight’ [pat<sup>3</sup>], ‘nine’ [kəu<sup>2</sup>], and ‘ten’ [sep<sup>2</sup>]), 17 bisyllabic, and 72 trisyllabic test CNW for the four speakers. The MNW ‘one’ [jɛt<sup>5</sup>] can occur only as a final syllable in the bisyllabic and trisyllabic CNW. The MNW from ‘two’ to ‘nine’ can occur as an initial and as a final syllable of the bisyllabic and trisyllabic CNW. As for the MNW ‘ten’, it can occur as an initial and as a final syllable of the bisyllabic CNW, but only as a medial syllable of the trisyllabic CNW. The numerals in the parentheses indicate the number of utterance, for instance, as shown in Table 1a for the MNW [sam<sup>55</sup>] ‘three’, the duration value of 351 ms is for a single utterance or token of the MNW [sam<sup>55</sup>] ‘three’. Similarly, the duration value of 221 ms is for a single token of the initial syllable of the bisyllabic CNW [sam<sup>55</sup> sep<sup>2</sup>] ‘three ten’, i.e., ‘thirty’; and the duration value of 311 ms is for a single token of the final syllable of the bisyllabic CNW [sep<sup>2</sup> sam<sup>55</sup>] ‘ten three’, i.e., ‘thirteen’. The duration value of 177 ms however is the mean value for 9 utterances of [sam<sup>55</sup>] ‘three’ which are the initial syllables of the trisyllabic CNW [sam<sup>55</sup> sep<sup>2</sup> jɛt<sup>5</sup>] ‘three ten one’, i.e., ‘thirty-one’, [sam<sup>55</sup> sep<sup>2</sup> ji<sup>22</sup>] ‘three ten two’, i.e., ‘thirty-two’, [sam<sup>55</sup> sep<sup>2</sup> sam<sup>55</sup>] ‘three ten three’, i.e., ‘thirty-three’ ... [sam<sup>55</sup> sep<sup>2</sup> kəu<sup>2</sup>] ‘three ten nine’, i.e., ‘thirty-nine’. Similarly, the duration value of 291 ms is the mean value for 8 utterances of [sam<sup>55</sup>] ‘three’ which are the final syllables of the trisyllabic CNW. And, as shown in the table, the duration value of 82 ms is the mean value for 72 utterances of the MNW [sep<sup>2</sup>] ‘ten’ which are all medial syllables of the 72 trisyllabic CNW. The dash lines indicate non-occurrence of a MNW.

As shown in the tables, for all the four speakers, (i) the rime durations of the component MNW of the bisyllabic and trisyllabic CNW are all shorter than those of the MNW in isolation; and (ii) the amount of reduction of the rime duration varies according to the position of the syllables in the CNW and the length of the CNW. The reduction of rime duration is greater for the non-final syllables than for the final syllables of the bisyllabic and in

particular trisyllabic CNW. In the bisyllabic CNW, the reduction of the rime duration for the initial syllable ranges from 22% to 59%. For the initial and medial syllables of the trisyllabic CNW, it ranges from 34% to 74% and from 33% to 47%, respectively. As for the final syllables of the CNW, in general the rime duration also reduces, but to a lesser degree, ranging from 1% to 33%.

In a few cases where the CNW-final MNW contain a short vowel and a stop ending, such as ‘one’ [jɛt<sup>2</sup>] and ‘ten’ [sɛp<sup>2</sup>], the rime duration for the MNW become slightly longer than that for the MNW in isolation, which is unexpected, e.g., for Male Speaker 1 (Table 1a), the rime for the MNW [jɛt<sup>2</sup>] ‘one’ is 137 ms, but 141 ms and 140 ms when the MNW is in the final position of a bisyllabic or trisyllabic CNW.

Tables 1a-1d: *The mean durations (in ms) of the rimes of the 10 component syllables, ‘one’ to ‘ten’, of the MNW and bisyllabic and trisyllabic CNW for the four speakers.*

(1a) Male Speaker 1

Component syllables	MNW	Bisyllabic CNW		Trisyllabic CNW		
		1 <sup>st</sup>	2 <sup>nd</sup>	1 <sup>st</sup>	2 <sup>nd</sup>	3 <sup>rd</sup>
[jɛt <sup>2</sup> ] ‘one’	137 (1)	-----	141 (1)	-----	-----	140 (8)
[ji <sup>22</sup> ] ‘two’	362 (1)	264 (1)	262 (1)	153 (9)	-----	278 (8)
[sam <sup>55</sup> ] ‘three’	351 (1)	221 (1)	311 (1)	177 (9)	-----	291 (8)
[sei <sup>33</sup> ] ‘four’	327 (1)	169 (1)	289 (1)	138 (9)	-----	281 (8)
[m <sup>23</sup> ] ‘five’	338 (1)	232 (1)	314 (1)	181 (9)	-----	270 (8)
[lɔk <sup>2</sup> ] ‘six’	120 (1)	64 (1)	104 (1)	61 (9)	-----	105 (8)
[ts <sup>h</sup> ɛt <sup>2</sup> ] ‘seven’	134 (1)	68 (1)	125 (1)	69 (9)	-----	122 (8)
[pat <sup>2</sup> ] ‘eight’	199 (1)	133 (1)	170 (1)	113 (9)	-----	177 (8)
[kɛu <sup>25</sup> ] ‘nine’	316 (1)	185 (1)	284 (1)	152 (9)	-----	279 (8)
[sɛp <sup>2</sup> ] ‘ten’	123 (1)	84 (9)	107 (8)	-----	82 (72)	-----

(1b) Male Speaker 2

Component syllables	MNW	Bisyllabic CNW		Trisyllabic CNW		
		1 <sup>st</sup>	2 <sup>nd</sup>	1 <sup>st</sup>	2 <sup>nd</sup>	3 <sup>rd</sup>
[jɛt <sup>2</sup> ] ‘one’	116 (1)	-----	119 (1)	-----	-----	117 (8)
[ji <sup>22</sup> ] ‘two’	296 (1)	161 (1)	281 (1)	127 (9)	-----	254 (8)
[sam <sup>55</sup> ] ‘three’	296 (1)	176 (1)	259 (1)	151 (9)	-----	244 (8)
[sei <sup>33</sup> ] ‘four’	278 (1)	143 (1)	235 (1)	111 (9)	-----	243 (8)
[m <sup>23</sup> ] ‘five’	326 (1)	178 (1)	279 (1)	140 (9)	-----	296 (8)
[lɔk <sup>2</sup> ] ‘six’	92 (1)	66 (1)	62 (1)	59 (9)	-----	86 (8)
[ts <sup>h</sup> ɛt <sup>2</sup> ] ‘seven’	92 (1)	72 (1)	82 (1)	61 (9)	-----	85 (8)
[pat <sup>2</sup> ] ‘eight’	218 (1)	120 (1)	205 (1)	103 (9)	-----	192 (8)
[kɛu <sup>25</sup> ] ‘nine’	290 (1)	158 (1)	280 (1)	136 (9)	-----	253 (8)
[sɛp <sup>2</sup> ] ‘ten’	110 (1)	65 (9)	97 (8)	-----	59 (72)	-----

(1c) Female Speaker 1

Component syllables	MNW	Bisyllabic CNW		Trisyllabic CNW		
		1 <sup>st</sup>	2 <sup>nd</sup>	1 <sup>st</sup>	2 <sup>nd</sup>	3 <sup>rd</sup>
[jɛt <sup>2</sup> ] ‘one’	172 (1)	-----	161 (1)	-----	-----	153 (8)
[ji <sup>22</sup> ] ‘two’	406 (1)	163 (1)	323 (1)	116 (9)	-----	348 (8)
[sam <sup>55</sup> ] ‘three’	386 (1)	220 (1)	319 (1)	142 (9)	-----	326 (8)
[sei <sup>33</sup> ] ‘four’	351 (1)	152 (1)	308 (1)	92 (9)	-----	323 (8)
[m <sup>23</sup> ] ‘five’	417 (1)	214 (1)	366 (1)	157 (9)	-----	377 (8)
[lɔk <sup>2</sup> ] ‘six’	154 (1)	88 (1)	121 (1)	66 (9)	-----	121 (8)
[ts <sup>h</sup> ɛt <sup>2</sup> ] ‘seven’	152 (1)	80 (1)	136 (1)	58 (9)	-----	122 (8)
[pat <sup>2</sup> ] ‘eight’	261 (1)	161 (1)	224 (1)	106 (9)	-----	219 (8)
[kɛu <sup>25</sup> ] ‘nine’	397 (1)	200 (1)	361 (1)	132 (9)	-----	350 (8)
[sɛp <sup>2</sup> ] ‘ten’	130 (1)	78 (9)	142 (8)	-----	69 (72)	-----

(1d) Female Speaker 2

Component syllables	MNW	Bisyllabic CNW		Trisyllabic CNW		
		1 <sup>st</sup>	2 <sup>nd</sup>	1 <sup>st</sup>	2 <sup>nd</sup>	3 <sup>rd</sup>
[jɛt <sup>2</sup> ] ‘one’	197 (1)	-----	156 (1)	-----	-----	164 (8)
[ji <sup>22</sup> ] ‘two’	422 (1)	212 (1)	339 (1)	146 (9)	-----	334 (8)
[sam <sup>55</sup> ] ‘three’	426 (1)	225 (1)	379 (1)	156 (9)	-----	352 (8)
[sei <sup>33</sup> ] ‘four’	422 (1)	173 (1)	337 (1)	120 (9)	-----	369 (8)
[m <sup>23</sup> ] ‘five’	459 (1)	216 (1)	331 (1)	151 (9)	-----	383 (8)
[lɔk <sup>2</sup> ] ‘six’	195 (1)	100 (1)	178 (1)	85 (9)	-----	180 (8)
[ts <sup>h</sup> ɛt <sup>2</sup> ] ‘seven’	183 (1)	84 (1)	154 (1)	77 (9)	-----	159 (8)
[pat <sup>2</sup> ] ‘eight’	272 (1)	151 (1)	269 (1)	126 (9)	-----	270 (8)
[kɛu <sup>25</sup> ] ‘nine’	440 (1)	181 (1)	399 (1)	144 (9)	-----	384 (8)
[sɛp <sup>2</sup> ] ‘ten’	144 (1)	90 (9)	147 (8)	-----	76 (72)	-----

### 3.1.2. Duration of initial consonant

Tables 2a-2d show the mean VOT and durations of the initial consonants which include the unaspirated stop [p] or [k], aspirated affricate [ts<sup>h</sup>], fricative [s], lateral approximant [l], and glide [j] of the ten basic component syllables of the MNW and of the bisyllabic and trisyllabic CNW for the four speakers. As will be shown in the following paragraphs, there is no general pattern of duration (including VOT) reduction for all the different initial consonants of the MNW in the context of the bisyllabic or trisyllabic CNW for the four speakers. In most cases, the duration is reduced or even deleted. However, there are a few cases in which the consonant duration increases.

Tables 2a-2d: *The mean durations (in ms) of the initial consonants of the 10 component syllables, ‘one’ to ‘ten’, of the MNW and bisyllabic and trisyllabic CNW for the four speakers.*

(2a) Male Speaker 1

Component syllables	MNW	Bisyllabic CNW		Trisyllabic CNW		
		1 <sup>st</sup>	2 <sup>nd</sup>	1 <sup>st</sup>	2 <sup>nd</sup>	3 <sup>rd</sup>
[jɛt <sup>2</sup> ] ‘one’	33 (1)	-----	53 (1)	-----	-----	53 (8)
[ji <sup>22</sup> ] ‘two’	9 (1)	0 (1)	40 (1)	6 (9)	-----	29 (8)
[sam <sup>55</sup> ] ‘three’	149 (1)	155 (1)	124 (1)	125 (9)	-----	125 (8)
[sei <sup>33</sup> ] ‘four’	163 (1)	147 (1)	142 (1)	139 (9)	-----	143 (8)
[m <sup>23</sup> ] ‘five’	0 (1)	0 (1)	0 (1)	0 (9)	-----	0 (8)
[lɔk <sup>2</sup> ] ‘six’	92 (1)	38 (1)	83 (1)	43 (9)	-----	84 (8)
[ts <sup>h</sup> ɛt <sup>2</sup> ] ‘seven’	119 (1)	102 (1)	93 (1)	83 (9)	-----	91 (8)
[pat <sup>2</sup> ] ‘eight’	7 (1)	11 (1)	9 (1)	10 (9)	-----	9 (8)
[kɛu <sup>25</sup> ] ‘nine’	15 (1)	13 (1)	14 (1)	17 (9)	-----	15 (8)
[sɛp <sup>2</sup> ] ‘ten’	176 (1)	135 (9)	119 (8)	-----	75 (72)	-----

(2b) Male Speaker 2

Component syllables	MNW	Bisyllabic CNW		Trisyllabic CNW		
		1 <sup>st</sup>	2 <sup>nd</sup>	1 <sup>st</sup>	2 <sup>nd</sup>	3 <sup>rd</sup>
[jɛt <sup>2</sup> ] ‘one’	14 (1)	-----	15 (1)	-----	-----	15 (8)
[ji <sup>22</sup> ] ‘two’	32 (1)	0 (1)	42 (1)	45 (9)	-----	18 (8)
[sam <sup>55</sup> ] ‘three’	141 (1)	128 (1)	84 (1)	139 (9)	-----	71 (8)
[sei <sup>33</sup> ] ‘four’	172 (1)	156 (1)	97 (1)	143 (9)	-----	80 (8)
[m <sup>23</sup> ] ‘five’	0 (1)	0 (1)	0 (1)	0 (9)	-----	0 (8)
[lɔk <sup>2</sup> ] ‘six’	74 (1)	33 (1)	63 (1)	39 (9)	-----	49 (8)
[ts <sup>h</sup> ɛt <sup>2</sup> ] ‘seven’	139 (1)	79 (1)	106 (1)	94 (9)	-----	101 (8)
[pat <sup>2</sup> ] ‘eight’	8 (1)	10 (1)	12 (1)	9 (9)	-----	10 (8)
[kɛu <sup>25</sup> ] ‘nine’	19 (1)	20 (1)	18 (1)	18 (9)	-----	18 (8)
[sɛp <sup>2</sup> ] ‘ten’	174 (1)	139 (9)	79 (8)	-----	66 (72)	-----

(2c) Female Speaker 1

Component syllables	MNW	Bisyllabic CNW		Trisyllabic CNW		
		1 <sup>st</sup>	2 <sup>nd</sup>	1 <sup>st</sup>	2 <sup>nd</sup>	3 <sup>rd</sup>
[jɛt <sup>2</sup> ] ‘one’	18 (1)	-----	37 (1)	-----	-----	28 (8)
[ji <sup>22</sup> ] ‘two’	29 (1)	31 (1)	28 (1)	20 (9)	-----	26 (8)
[sam <sup>55</sup> ] ‘three’	132 (1)	136 (1)	118 (1)	101 (9)	-----	116 (8)
[sei <sup>33</sup> ] ‘four’	165 (1)	130 (1)	135 (1)	106 (9)	-----	127 (8)
[m <sup>23</sup> ] ‘five’	17 (1)	20 (1)	30 (1)	0 (9)	-----	0 (8)
[lɔk <sup>2</sup> ] ‘six’	89 (1)	40 (1)	87 (1)	34 (9)	-----	71 (8)
[ts <sup>h</sup> ɛt <sup>2</sup> ] ‘seven’	92 (1)	92 (1)	98 (1)	65 (9)	-----	94 (8)
[pat <sup>2</sup> ] ‘eight’	7 (1)	8 (1)	8 (1)	8 (9)	-----	8 (8)
[kɛu <sup>25</sup> ] ‘nine’	15 (1)	11 (1)	12 (1)	14 (9)	-----	15 (8)
[sɛp <sup>2</sup> ] ‘ten’	166 (1)	123 (9)	103 (8)	-----	71 (72)	-----

(2d) Female Speaker 2

Component syllables	MNW	Bisyllabic CNW		Trisyllabic CNW		
		1 <sup>st</sup>	2 <sup>nd</sup>	1 <sup>st</sup>	2 <sup>nd</sup>	3 <sup>rd</sup>
[jɛt <sup>2</sup> ] ‘one’	18 (1)	-----	42 (1)	-----	-----	34 (8)
[ji <sup>22</sup> ] ‘two’	36 (1)	0 (1)	40 (1)	0 (9)	-----	34 (8)
[sam <sup>55</sup> ] ‘three’	146 (1)	157 (1)	110 (1)	109 (9)	-----	93 (8)
[sei <sup>33</sup> ] ‘four’	156 (1)	145 (1)	112 (1)	121 (9)	-----	97 (8)
[m <sup>23</sup> ] ‘five’	0 (1)	0 (1)	75 (1)	0 (9)	-----	0 (8)
[lɔk <sup>2</sup> ] ‘six’	97 (1)	36 (1)	64 (1)	56 (9)	-----	60 (8)
[ts <sup>h</sup> ɛt <sup>2</sup> ] ‘seven’	98 (1)	81 (1)	80 (1)	68 (9)	-----	69 (8)
[pat <sup>2</sup> ] ‘eight’	12 (1)	9 (1)	12 (1)	9 (9)	-----	10 (8)
[kɛu <sup>25</sup> ] ‘nine’	16 (1)	14 (1)	15 (1)	16 (9)	-----	17 (8)
[sɛp <sup>2</sup> ] ‘ten’	156 (1)	158 (9)	100 (8)	-----	66 (72)	-----

As shown in Tables 2a-2d, for all the four speakers, the VOT of the initial unaspirated stop of the MNW [pat<sup>3</sup>] ‘eight’ or [kəu<sup>25</sup>] ‘nine’ may reduce or increase when the MNW occurs in the bisyllabic and trisyllabic CNW. There is no general pattern of reduction or increase for the four speakers, and the change in either direction is small. As for the VOT of the initial aspirated affricate [ts<sup>h</sup>] of the MNW [ts<sup>h</sup>et<sup>5</sup>] ‘seven’, it tends to shorten when the MNW occurs in the bisyllabic or trisyllabic CNW. This is true for the two male speakers and Female Speaker 2. For Female Speaker 1, there is a slight increase when the MNW occurs in the final position of the CNW, bisyllabic or trisyllabic.

As for the durations of the initial glide [j] of the MNW [jet<sup>5</sup>] ‘one’ and [ji<sup>22</sup>] ‘two’, they may increase or reduce when the MNW occurs in the bisyllabic and trisyllabic CNW and again there is no pattern of change in either direction. In some cases in the speech of the two male speakers and Female Speaker 2, the initial [j] of [ji<sup>22</sup>] ‘two’ does not occur in the CNW-initial position. This may be due to the fact that in HKC the occurrence of [j] of a monosyllable is optional when it occurs before a high front vowel [i] ([3]).

As for the duration of the initial lateral approximant [l] of the MNW [luk<sup>2</sup>] ‘six’, it reduces in the contexts of the bisyllabic and trisyllabic CNW, especially in the CNW-initial position, for all the four speakers. In the initial position of the bisyllabic or trisyllabic CNW, the duration reduction of the initial [l] ranges from 45% to 60%, and when the MNW occurs in the CNW-final position, the duration reduction ranges from 10% to 30%.

The durations of the initial fricative [s] of the MNW [sam<sup>55</sup>] ‘three’, [sei<sup>33</sup>] ‘four’, and [sep<sup>2</sup>] ‘ten’ in most cases reduce when the MNW occurs in the bisyllabic and trisyllabic CNW. In four of the bisyllabic CNW, the duration of the fricative for the MNW in the CNW-initial position increases slightly. The degree of duration reduction of the fricative for the MNW varies according to the position in which the MNW occurs. The largest reduction takes place when the MNW occurs in the medial position in the trisyllabic CNW, followed by the final position in the trisyllabic CNW and then the final position in the bisyllabic CNW. The least reduction for [s] is when the MNW occurs in the CNW-initial position.

### 3.1.3. Closure periods of the bisyllabic and trisyllabic CNW

Tables 3a-3b show the mean durations of the closures in the bisyllabic and trisyllabic CNW which contain a sequence of an MNW [sep<sup>2</sup>] ‘ten’ followed by one of the other nine MNW from ‘one’ to ‘nine’, e.g., [sep<sup>2</sup> jet<sup>5</sup>] ‘ten one’, i.e., ‘eleven’; [sep<sup>2</sup> ji<sup>22</sup>] ‘ten two’, i.e., ‘twelve’ ... [sep<sup>2</sup> kəu<sup>25</sup>] ‘ten nine’, i.e., ‘nineteen’; [ji<sup>22</sup> sep<sup>2</sup> jet<sup>5</sup>] ‘two ten one’, i.e., ‘twenty-one’; [ji<sup>22</sup> sep<sup>2</sup> ji<sup>22</sup>] ‘two ten two’, i.e., ‘twenty-two’ ... [ji<sup>22</sup> sep<sup>2</sup> kəu<sup>25</sup>] ‘two ten nine’, i.e., ‘twenty-nine’; [sam<sup>55</sup> sep<sup>2</sup> jet<sup>5</sup>] ‘three ten one’, i.e., ‘thirty-one’; [sam<sup>55</sup> sep<sup>2</sup> ji<sup>22</sup>] ‘three ten two’, i.e., ‘thirty-two’ ... [kəu<sup>25</sup> sep<sup>2</sup> kəu<sup>25</sup>] ‘nine ten nine’, i.e., ‘ninety-nine’. For each speaker, the sequence of [sep<sup>2</sup>] ‘ten’ followed by one of the nine MNW occurs only once in the bisyllabic CNW, and eight times in the trisyllabic CNW. The number of occurrence is indicated by the numeral in the parentheses. Tables 3a and 3b show the mean durations of closures for the two males speakers and two female speakers, respectively.

A few observations can be made about the closure durations. Firstly, as shown in Table 3a for the two male speakers, the closure durations for the bisyllabic CNW are longer than the durations in the same phonetic contexts for the trisyllabic CNW, e.g., for Male Speaker 1 the closure durations in [sep<sup>2</sup> jet<sup>5</sup>] ‘ten one’ are 117 ms for the bisyllabic CNW and 82 ms for the trisyllabic CNW; and for Male Speaker 2 they are 83 ms for the bisyllabic CNW and 49 ms for the trisyllabic CNW. This is also true for the two female speakers (Table 3b). Secondly, as shown in Tables 3a and 3b, for all the four speakers, the closure durations

are the shortest when [sep<sup>2</sup>] ‘ten’ is followed (i) by the alveolar fricative [s] of [sam<sup>55</sup>] ‘three’ or [sei<sup>33</sup>] ‘four’, (ii) by the alveolar lateral [l] of [luk<sup>2</sup>] ‘six’, or (iii) by the syllabic bilabial nasal [m] of [m<sup>23</sup>] ‘five’. However, there are two exceptions, the closure duration before [m] of [m<sup>23</sup>] for Male Speaker 1 and the closure duration before [l] of [luk<sup>2</sup>] for Female Speaker 1 tend to be longer. Thirdly, the closure durations before the palatal glide [j] of [jet<sup>5</sup>] ‘one’ and [ji<sup>22</sup>] ‘two’ tend to be longer than the closure durations before [s l m]. And finally, the closure durations before the obstruents [ts<sup>h</sup> p k] tend to be the longest.

Tables 3a-3b: *The mean durations (in ms) of the closure periods in the bisyllabic and trisyllabic CNW containing the MNW ‘ten’ followed by one of the 9 MNW from ‘one’ to ‘nine’ for the four speakers (‘B’= bisyllabic; ‘T’= trisyllabic).*

(3a) Male Speakers

Sequences of [sep <sup>2</sup> ] ‘ten’ followed by another MNW	Male Speaker 1		Male Speaker 2	
	B. CNW	T. CNW	B. CNW	T. CNW
[sep <sup>2</sup> jet <sup>5</sup> ] ‘ten one’	117 (1)	82 (8)	83 (1)	49 (8)
[sep <sup>2</sup> ji <sup>22</sup> ] ‘ten two’	123 (1)	62 (8)	68 (1)	52 (8)
[sep <sup>2</sup> sam <sup>55</sup> ] ‘ten three’	83 (1)	47 (8)	52 (1)	38 (8)
[sep <sup>2</sup> sei <sup>33</sup> ] ‘ten four’	51 (1)	40 (8)	57 (1)	33 (8)
[sep <sup>2</sup> m <sup>23</sup> ] ‘ten five’	167 (1)	99 (8)	66 (1)	47 (8)
[sep <sup>2</sup> luk <sup>2</sup> ] ‘ten six’	63 (1)	50 (8)	51 (1)	38 (8)
[sep <sup>2</sup> ts <sup>h</sup> et <sup>5</sup> ] ‘ten seven’	120 (1)	91 (8)	81 (1)	48 (8)
[sep <sup>2</sup> pat <sup>3</sup> ] ‘ten eight’	167 (1)	138 (8)	84 (1)	71 (8)
[sep <sup>2</sup> kəu <sup>25</sup> ] ‘ten nine’	182 (1)	138 (8)	98 (1)	76 (8)

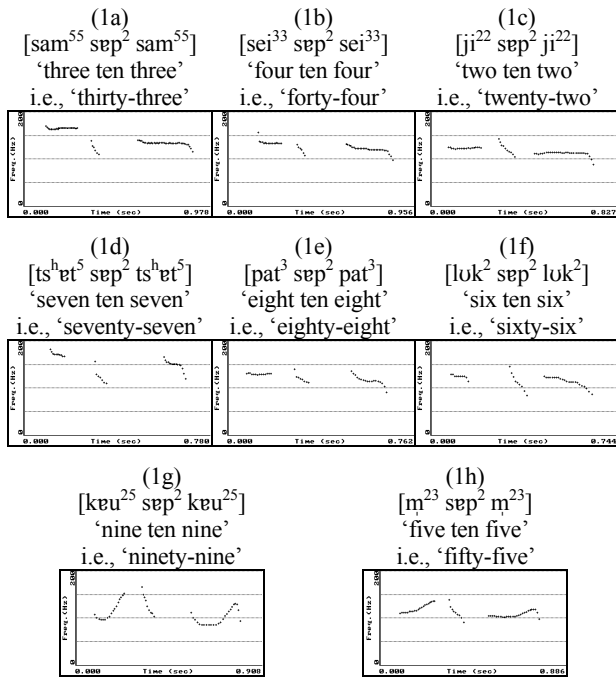
(3b) Female Speakers

Sequences of [sep <sup>2</sup> ] ‘ten’ followed by another MNW	Female Speaker 1		Female Speaker 2	
	B. CNW	T. CNW	B. CNW	T. CNW
[sep <sup>2</sup> jet <sup>5</sup> ] ‘ten one’	102 (1)	82 (8)	87 (1)	58 (8)
[sep <sup>2</sup> ji <sup>22</sup> ] ‘ten two’	150 (1)	83 (8)	80 (1)	41 (8)
[sep <sup>2</sup> sam <sup>55</sup> ] ‘ten three’	52 (1)	35 (8)	72 (1)	28 (8)
[sep <sup>2</sup> sei <sup>33</sup> ] ‘ten four’	47 (1)	32 (8)	76 (1)	28 (8)
[sep <sup>2</sup> m <sup>23</sup> ] ‘ten five’	72 (1)	56 (8)	80 (1)	42 (8)
[sep <sup>2</sup> luk <sup>2</sup> ] ‘ten six’	104 (1)	75 (8)	79 (1)	47 (8)
[sep <sup>2</sup> ts <sup>h</sup> et <sup>5</sup> ] ‘ten seven’	99 (1)	85 (8)	108 (1)	63 (8)
[sep <sup>2</sup> pat <sup>3</sup> ] ‘ten eight’	126 (1)	109 (8)	96 (1)	66 (8)
[sep <sup>2</sup> kəu <sup>25</sup> ] ‘ten nine’	145 (1)	115 (8)	132 (1)	91 (8)

As for the sequences of [ts<sup>h</sup>et<sup>5</sup> sep<sup>2</sup>] ‘seven ten’ and [pat<sup>3</sup> sep<sup>2</sup>] ‘eight ten’, in which the initial MNW ends with an alveolar [t], no closure period is found between the two successive syllables. This indicates that the final stop ending [t] is deleted when it is followed by a homorganic alveolar fricative [s]. However, the final stop ending [k] is retained when followed by an alveolar fricative [s] in the sequence of [luk<sup>2</sup> sep<sup>2</sup>] ‘six ten’, as closure period is found in the sequence. The closure durations after [k] in the bisyllabic CNW are 63 ms (Male Speaker 1), 49 ms (Male Speaker 2), 40 ms (Female Speaker 1), and 34 ms (Female Speaker 2). The mean closure durations after [k] in the trisyllabic CNW are 36 ms (Male Speaker 1), 33 ms (Male Speaker 2), 28 ms (Female Speaker 1), and 20 ms (Female Speaker 2). Again, the closure duration after the final [k] is shorter in the trisyllabic CNW than in the bisyllabic CNW, similar to the patterns of the closure duration after [p] of [sep<sup>2</sup>] ‘ten’ shown in Tables 3a-3b.

### 3.2. Fundamental frequency (F<sub>0</sub>)

The citation forms of the tones on the 10 MNW from ‘one’ to ‘ten’ are maintained generally, when the MNW occur in the bisyllabic and trisyllabic CNW. While the shape of the F<sub>0</sub> contour for any MNW in the CNW is retained, the F<sub>0</sub> level may change, depending on the position of the MNW in the CNW. Figures 1a-1h and 2a-2h show the F<sub>0</sub> contours of the trisyllabic CNW, with the first and final MNW being associated with the same citation tone, [55], [33], [22], [5], [3], [2], [25], or [23] in HKC for Male Speaker 1 and Female Speaker 1. Due to page limit, the F<sub>0</sub> contours of the other two speakers are not presented here.

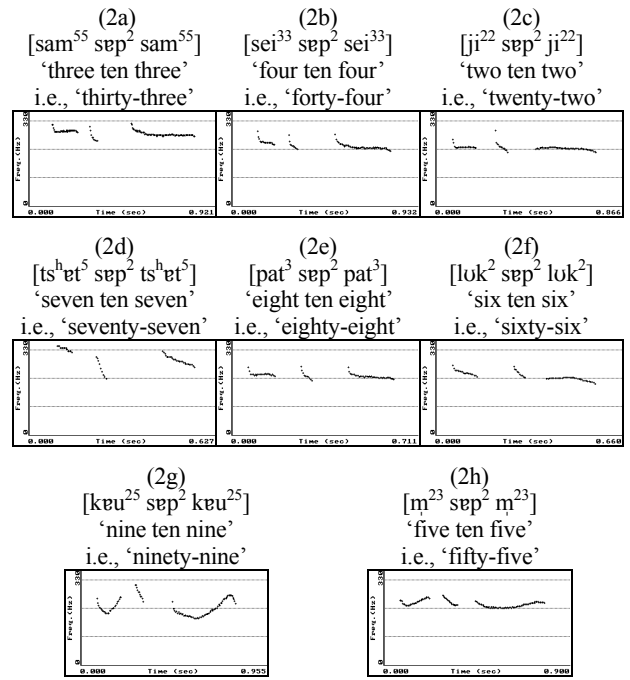


Figures 1a-1h: The  $F_0$  contours for the trisyllabic CNW with the same citation tone [55], [33], [22], [5], [3], [2], [25], or [23] on the CNW-initial and CNW-final MNW for Male Speaker 1.

As shown in Figures 1a-1c and Figures 2a-2c, for both the male and female speakers the  $F_0$  contours of the citation tones [55 33 22] on the CNW-initial and CNW-final MNW of the trisyllabic CNW are maintained, i.e., they are basically level. However, the  $F_0$  level of each of the citation tones is slightly lowered for the CNW-final MNW, which is due probably to the effect of  $F_0$  declination. The  $F_0$  level lowering is also true for the  $F_0$  contours of the short citation tones [5 3 2] (Figures 1d-1f and 2d-2f), i.e., the slightly falling  $F_0$  contours of [5 3 2] are maintained in the trisyllabic CNW and the  $F_0$  levels of these tones are lowered in the CNW-final position. Similarly, the  $F_0$  contours of the rising citation tones [25 23] (Figures 1g-1h and 2g-2h) are maintained and the levels of the  $F_0$  contours of these tones are lowered for the MNW in the CNW-final position. In all the cases (Figures 1a-1h and 2a-2h), the tonal durations for the MNW in the initial position of the CNW are shorter than those for the MNW in the final position. The tonal patterns for the CNW for Male Speaker 2 and Female Speaker 2 which are not presented here are similarly to those for Male Speaker 1 and Female Speaker 1 which have just been presented.

### 3.3. Intensity

The intensity levels of the component MNW of the bisyllabic or trisyllabic CNW vary, depending on the position of the component MNW in the CNW as well as the tones on the component MNW. For those MNW which contain the same rime and are associated with the same citation tone, the intensity levels tend to be slightly lower for the component MNW in the bisyllabic or trisyllabic CNW-final position than in the non-final position. This correlates positively with the  $F_0$  level difference between the final and non-final MNW of the CNW. However, for those MNW which are associated with a rising tone [25] or [23], the  $F_0$  contours and the intensity curves are negatively correlated, that is, while the  $F_0$  contour rises, the intensity curve falls. Due to the page limit, the intensity curves of the bisyllabic and trisyllabic CNW for the four speakers are not presented here.



Figures 2a-2h: The  $F_0$  contours for the trisyllabic CNW with the same citation tone [55], [33], [22], [5], [3], [2], [25], or [23] on the CNW-initial and CNW-final MNW for Female Speaker 1.

## 4. Conclusion

The paper has presented the results of the acoustic analysis of the prosodic characteristics, i.e., duration,  $F_0$ , and intensity, of the MNW and the component MNW of the bisyllabic and trisyllabic CNW in HKC. The prosodic characteristics vary according to the position in which the component MNW occur in the CNW. The intra-speaker and cross-speaker variability is small in the rime duration,  $F_0$  contour, or intensity curve for the CNW, but large in the duration of the syllable-initial consonants. The main difference in terms of prosody between the MNW in isolation and the component MNW of the bisyllabic and trisyllabic CNW is in duration. Thus, it is suggested that in synthesizing the CNW in HKC by way of concatenating a set of MNW, special attention should be paid to the temporal characteristics of the component MNW of the CNW.

## 5. References

- [1] Lee, W.S.; Chen, F.X.; Luke, K.K.; Shen, L.Q., 2002. The prosody of bisyllabic and polysyllabic words in Hong Kong Cantonese. *Proceedings of the 1<sup>st</sup> International Conference on Speech Prosody*, 451-454. Aix-en-Provence: Université de Provence.
- [2] Lee, W.S., 2003. A phonetic study of reduplicated syllables in Hong Kong Cantonese. *Proceedings of the 6<sup>th</sup> National Conference on Modern Phonetics*, (1)62-67. Tianjin: Tianjin Normal University.
- [3] Lee, W.S. (ms). Sound Pattern of Hong Kong Cantonese.
- [4] Lin, M.C.; Yan, J.Z.; Sun, G.H. (林茂燦, 顏景助, 孫國華), 1984. 北京話兩字組正常重音的初步實驗 (A preliminary experiment on the stress in the bisyllabic compounds in Beijing dialect). *Fangyan*(1), 57-73.
- [5] Yan, J.Z.; Lin, M.C. (顏景助, 林茂燦), 1988. 北京話三字組重音的聲學表現 (The acoustic manifestation of the stress in the trisyllabic compounds in Beijing dialect). *Fangyan*(3), 227-237.