Interactions between Syllabic-Level Prosody and Prosodic-Group Boundary Markers in Cantonese

Ivan Chow
The University of Western Ontario

Pitch reset (Swerts 1997, Swerts & Geluykens 1994), pauses (Shen 1992) and pre-boundary lengthening (Delais-Roussarie 1995, Gussenhoven & Rietveld 1992, Wightman *et al.* 1992) are found to be common juncture markers in the conveyance of prosodic grouping. This article discusses the interactions between syllabic-level prosody and prosodic juncture markers. In an acoustic study, we investigate the prosodic devices used by Cantonese speakers to mark junctures within the sentence. As relative pitch (F_0) height and syllabic duration play a contrastive role in Cantonese syllabic recognition, these prosodic juncture markers create potential interferences to syllable-level prosody. We hypothesized that (i) in the melodic dimension, pitch reset interacts with level and checked tones, creating potential tonal misrecognitions in perception. (ii) in the temporal dimension, lengthened pre-boundary syllables may interfere with the recognition of syllables which are distinguished by syllabic duration. Comparative data analyses revealed that, given the right circumstances, temporal and melodic variations caused by prosodic juncture marking can be large enough to cause miscomprehension. Moreover, individual differences in terms of the amplitude of these variations can be quite large, representing a good possibility of tone misrecognition due to melodic interference. When miscomprehension does occur, it provides us with a unique environment in which to study the impact of prosodic juncture markers in speech production and perception.

In an investigation of the Cantonese tone system, Vance recognizes the importance of the relation between syntactic structure and intonation: "Ladefoged (cited in Vance 1976:368) observed that both tonal and non-tonal languages use grammatical intonation (of which pitch reset is a component) to convey syntactic information." Given that fundamental frequency is the acoustic correlate of both intonation and lexical tone, one may wonder to what extent syntactically-motivated intonation patterns may affect the contours of Cantonese tones, and vice versa (Bauer & Benedict 1997:148). As most studies on the relation between lexical tones and intonation have been conducted with Mandarin as a subject language (e.g. Chao 1930, Zee 1993), (with the majority concentrating on the production and perception of third tone sandhi (c.f. Shih 1986, Xu 1997, Zhang 1997), less is known about the interactions between tone and intonation in Cantonese (cf. Gandour 1981, Lee 2004). Mandarin has four lexical tones that are distinguished by both shape and relative pitch, whereas the Hong Kong variety of Cantonese has nine lexical tones. Among the nine tones, up to six are distinguished only by relative pitch (see section 2 below). If a prosodic boundary happens to be located between two lexical tones that are

Actes du congrès annuel de l'Association canadienne de linguistique 2006.
Proceedings of the 2006 annual conference of the Canadian Linguistic Association.

only distinguished by relative pitch height, the resetting of pitch height may coincide with the pitch height difference between lexical tones, thereby creating a situation in which tone recognition may be severely hindered.

In the temporal dimension, syllabic duration is a contrastive feature in Cantonese syllable recognition (Kao 1971:39-40, Lee 2004, Li 1985:21-31, Zee 1993). For example, in the recognition of the long diphthong [a:j] vs. the short [ɐj], despite the fact that the two vowels are transcribed with different phonetic symbols, indicating difference in vowel quality, the primary feature that distinguishes between the two rimes is the difference in vowel length rather than quality (Barrie 2003; Zee 1993:98, c.f.Bauer & Benedict 1997). On the other hand, Chow (2005b) observed that lengthened syllables can be up to 9% longer than their non-lengthened counter-parts. As a result, questions arise as to whether pre-boundary lengthening is prominent enough to trigger potential syllabic misrecognition?

Data from Chow's acoustic analyses (2005b) are examined in detail. Based on available acoustic data on Cantonese citation tones (Bauer & Benedict 1997, Gandour 1981, Kwok & Luk 1986, Lee 2004, Vance 1976), this article discusses potential interactions between these juncture markers and syllabic-level prosody in the melodic and temporal dimensions. It is our goal to give an overview of the phenomena in question, and to generate constructive suggestions from other researchers in order to fine-tune our experimental design and methodology.

In section 1, we begin with a brief introduction to the Cantonese tone system. We then examine the melodic interactions between pitch reset and level tones by comparing measured pitch height difference and the relative pitch height of the citation tones in order to facilitate our discussion regarding the potential perceptual interference of pitch reset towards tone recognition. Section 2 discusses interactions in the temporal dimension between syllabic duration and pre-boundary lengthening. In this section, we explore the possibility of tone/morpheme misrecognition by comparing our measurements of syllable lengthening to the syllabic duration of citation tones. Concluding remarks are found in section 3.

1. Melodic interactions

1.1. Cantonese lexical tones

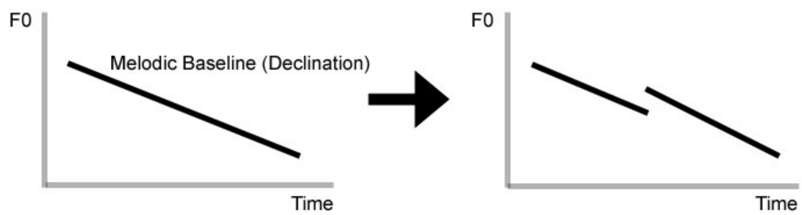
There are nine citation tones in the variety of Cantonese spoken in Hong Kong (55, 33, 22, 21[1], 25, 23, 5, 3, and 2). 25 and 23 are rising tones; 55, 33, 22 and 21 are level tones. Tones 5, 3 and 2 are level tones with shorter syllabic duration, referred to as checked tones. These tones associate with syllables ending with consonantal stops. Rising tones differ in both pitch register and

---

[1] The falling tone (21) often surfaces as a low level tone in continuous speech (Bauer & Benedict 1997, Lee 2004, Mai 1998).

contour, whereas level tones and checked tones can only be distinguished by relative pitch height. A morpheme in Cantonese consists of the combination of a syllable and a lexical tone. It functions as both a phonological and morphological unit of analysis (Chan 1999). Melodic contours at the level of the intonational phrase are largely defined by the concatenation of lexical tones at the syllabic level (e.g. Chao 1930); a down-drifting melodic baseline (also referred to as declination, cf. Gussenhoven 1988) is then imposed upon this concatenated contour. Focus and sentence modality are conveyed by widespread melodic variations while keeping the shape of lexical tones more or less intact (Chow 2005a, Xu 1997).

1.2. Pitch reset



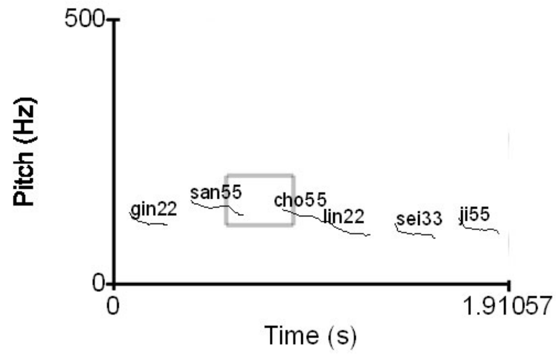
(1) Measurement of Pitch Reset

As shown in (1), pitch reset refers to the interruption of the otherwise continuously declining melodic baseline at the boundary location. After the boundary, pitch height is readjusted to a higher value where it continues to decline towards the end of the sentence / intonational phrase.

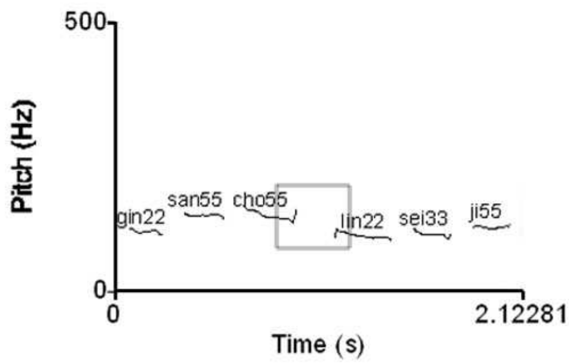
In a pilot study using syntactic disambiguation tasks, pitch reset was shown to be a prominent prosodic juncture marker in Cantonese (Chow 2005b). Figures 2 and 3 below show the measured melodic contours of a pair of test sentences. In line with Vance's (1976) observation concerning the relation between syntax and intonation, the stimuli are designed in such a way that each pair of sentences consists of an identical array of morphemes. Two different interpretations can be derived by conveying prosodic boundaries at different syntactically motivated locations. To compensate for the inherent pitch differences between different combinations of lexical tones, pitch height difference was measured across boundary locations (marked by squares). The measured pitch height difference is then compared to that of the same pair of syllables in the corresponding sentence where no boundary was present at that particular location. In figures 2 and 3, pitch reset is marked by opaque boxes. For example, the measured pitch height difference for the boundary between "san55" and "cho55" in figure 2 is compared to the pitch height difference between "cho55" and "lin55" in figure 3. Pitch height is measured within the nucleus portion of the syllable at the point of optimal intensity. Pitch height difference is then derived by subtracting the pitch height value of the pre-boundary syllable from that of the post-boundary syllable. The pitch height difference is then compared between the sentence *with* the boundary and the

Comment: Are these the correct morphemes? Should they be cho55 and lin22?

corresponding sentence *without* the boundary. A positive boundary-non-boundary difference signals an upward shift, whereas a negative difference signals a downward shift.



(2) Pitch contour for test sentence (a).
"gin22 san55 / cho55 lin22 sei33 ji55"
"Weight-lifting strengthens the limbs."



(3) Pitch contour for test sentence (b).
"gin22 san55 cho55 / lin22 sei 33 ji55"
"Aerobics exercises the limbs."

Using this experimental design, contextual variations between lexical tones are kept to a minimum, measured pitch height difference can be readily attributed to pitch reset. On average, an upward shift of 15.4Hz was found in 30 sentence samples. Considering the standard deviation, upward adjustment is likely going to be as high as 28Hz.

1.3. Interaction between pitch reset and inherent pitch height differences between level tones

Only a handful of acoustic studies concerning the perceptual aspect of Cantonese intonation have been published (Bauer & Benedict 1997, Gandour 1981, Kwok & Luk 1986, Lee 2004, Vance 1976). Bauer & Benedict observe that "tone contours tend to become progressively lower as the speaker arrives at the end of the utterance, the same citation tone will appear to have a lower pitch register than the one appearing earlier in the same sentence" (p.154). Lee conducted an acoustic study to examine the melodic manifestation of Cantonese citation tones in the declarative and interrogative contexts. Her experiment is designed in such a way that the citation tone under scrutiny is always in the focused position. Her results showed that in interrogative sentences, differences in pitch height between tones appearing at different locations of a sentence are reduced. In other words, the effects of declination is reduced in the interrogative mode compared to that of declarative sentences. In some cases, pitch height differences have disappeared altogether. As such, this observation clearly supports the view that melodic variations within larger phonological domains *does* influence the F_0 patterns of lexical tones.

Let us examine the data regarding pitch reset and relative pitch height as a distinguishing feature between lexical tones. Bauer & Benedict (1997) measure the F0 patterns of all Cantonese lexical tones produced in isolation by six native speakers of Cantonese. For comparison purposes, data from two native speakers of the Guangzhou variety were excluded from the present analysis. As a result, the subject pool consists of four native speakers of the Hong Kong variety of Cantonese, two male and two female.

	High-Level (55)	**Mid-Level (33)**	**Mid-Low-Level (22)**	**Mid-Low-Falling (21)**
Avg. F_0 (in Hz)	243.1	203.2	175.4	155.1
F_0 diff.	39.9	27.8	20.3	

(4) Inherent F_0 Difference between Level Tones

	High Checked (5)	**Mid-Checked (3)**	**Mid-Low-Checked (2)**
Avg. F_0 (in Hz)	242.8	198.8	174.7
F_0 diff.	44.0	24.1	

(5) Inherent F_0 Difference between Checked Tones

(4) and (5) show the average pitch heights of four level tones and three checked tones in Cantonese. The bottom row of each figure shows the calculated pitch height difference between two closest tones, e.g. the pitch height difference between high-level tone and its next lower tone, the mid-level is 39.9Hz. In

Bauer & Benedict's study, F_0 was measured at the onset, peak and end-point of the tone. Since we are concerned with difference in relative pitch height only, average F_0 values of the three measuring points were used for comparison in this study.

Deducing from Bauer & Benedict's measurements, F_0 differences between level and checked tones range from 20.3Hz to 44.0Hz. Compared to the data from Chow (2005b), F_0 differences between level tones are larger than the 15.4Hz average pitch adjustment caused by pitch reset. In normal circumstances, given a prosodic boundary between two level tones (e.g. a pre-boundary mid-low falling tone of 175.4Hz and a post-boundary mid-low tone of 155.1), pitch reset triggers an upward adjustment of 15.4Hz, increasing the pitch of the mid-low tone to 170.5Hz. Hence, the new (higher) F_0 of the mid-low tone is less than 5Hz apart from the pre-boundary mid-low falling tone. Moreover, as mentioned above, upward F_0 adjustments triggered by pitch reset can increase the F_0 value of a post-boundary tone by up to 28Hz (one standard deviation above average). In the recorded corpus collected in his study, F_0 values of the post-boundary tone were raised by as much as 60Hz in some cases, far exceeding the inherent F_0 difference between the any adjacent level tones. In fact, pitch adjustment exceeded 20Hz in 6 out of 33 instances, which comprises 18.2% of all cases. Such a large percentage is certainly not negligible. As such, it is entirely possible that pitch height adjustment due to pitch reset can cause tonal misrecognition in normal circumstances. However, as in many Chinese dialects, there are numerous restrictions in Cantonese in terms of possible combinations of syllable and tone, onset and rhyme, as well as onset and coda (Barrie 2003). Certain tone-syllable combinations are disallowed in the Cantonese syllabary. These restrictions may very well serve to reduce the chances of tonal misrecognition. Yet, when tonal misrecognition does occur, it provides a unique experimental condition in which to examine the interaction between tone and intonation in speech perception.

2. Temporal interactions

2.1. Cantonese Syllabic Structures subject to Durational Contrast

According to Bauer & Benedict (1997:409), regardless of lexical tone, the Cantonese syllabary recognizes 750 different combinations of initials and rhymes. Among the six types of rhymes listed in figure 6 below, syllables ending with the same coda consonant contrast in terms of vowel length (i.e. between type 1 and 2; 3 and 4; 5 and 6). Based on Kao (1971) and Li (1985), average vowel durations of these syllables are shown below: long vowels appear to be approximately twice as long as short vowels.

	Nucleus + Coda		
1	$V{:}C_{SV}$	Vowel (long) + semivowel	203ms
2	VC_{SV}	Vowel (short) + semivowel	100ms
3	$V{:}C_{nas}$	Vowel (long) + nasal	195ms
4	VC_{nas}	Vowel (short) + nasal	99ms
5	$V{:}C_{stop}$	Vowel (long) + stop	159ms
6	VC_{stop}	Vowel (short) + stop	83ms

(6) Rhymes in the Cantonese Syllables (Bauer & Benedict 1997:13, 38)

"Although most phonological studies have generally not treated vowel length as contrastive (except with respect to the low vowels), Bauer and Benedict (1997: 39) make use of duration measurements from Kao and Li to draw the conclusion that vowel duration should be considered a distinctive characteristic of the vowel system of Cantonese" (Chan 1999: 102). Some examples of minimal pairs created by contrasting vowel lengths are show in figure 7.

$V{:}C_{SV}$ vs. VC_{SV}	懷 [wa:j] 'to bear'	圍 [wɐj] 'to surround'
$V{:}C_{nas}$ vs. VC_{nas}	還 [wa:n] 'to return'	暈 [wɐn] 'to faint'
$V{:}C_{stop}$ vs. VC_{stop}	達 [da:t] "to achieve"	突 [dɐt] "sudden"

(7) Minimal Pairs Contrasted by Vowel Length

2.2. Pre-boundary Lengthening

As a prosodic boundary marker, pre-boundary lengthening refers to the phenomenon in which duration of the syllable immediately preceding a juncture is increased compared to syllables in the rest of the sentence. Shen (1992) ask native speakers of Mandarin to read a series of syntactically ambiguous sentences aloud while using prosody to clarify the utterance structure. Results from her study showed that in speech production, pauses were by far more effective than pre-boundary lengthening as juncture markers. Pre-boundary syllables would have to be lengthened by a significant amount for it to be worthwhile as a juncture marker. Such results are consistent with Chow's results (2005b). In his study, pre-boundary lengthening was found to be less frequent than the other markers under scrutiny—it was detected in only 63% of the test boundaries, as opposed to 98% for pauses and 88% for pitch reset. In addition, in the cases where pre-boundary lengthening was detected, the pre-boundary syllable was on average 1.8% longer than its non-boundary counterpart, whereas the observed values ranged from 1 to 9%.

2.3. Interaction between pre-boundary lengthening and contrastive syllabic duration

Bauer & Benedict's data indicate that Cantonese long syllables are on average twice as long as short syllables of the same syllabic (rhyme) structure. Data obtained from Chow's study indicate that lengthened pre-boundary syllables are only 1.8% (9% in the extreme case) longer. Compared to the inherent difference in syllabic durations between long and short syllables, the effects of pre-boundary lengthening are clearly not prominent enough to trigger syllabic misrecognition. Alternatively, compared to the results of similar studies in other languages (e.g. Delais-Roussarie (1995) for French, Swerts & Rietveld (1992) for Dutch and Wightman *et al.* (1992) for English), the effect of pre-boundary lengthening in Cantonese is comparatively smaller (up to 10% in Cantonese vs. 40% in French reported in Delais-Roussarie (1995)). Perhaps the fact that syllabic duration is a contrastive feature contributes in part to the speakers' tendency to limit their lengthening of pre-boundary syllables in order to avoid syllabic miscomprehension— a manner in which the Cantonese phonological system evolves to remain an effective means of communication. Nonetheless, this result provides another piece to the puzzle regarding the effects of language-specific experience in the processing of prosodic information.

3. Conclusions

Although we have shown that prosodic juncture markers *do* interact with syllabic duration and tonal features, not all types of juncture markers are likely to trigger miscomprehension. We looked at the impact of pitch reset towards level and checked tone recognition, as well as the interaction between pre-boundary lengthening and syllabic duration in certain types of Cantonese syllables. Average F_0 upward shifts caused by pitch reset are found to be close to the inherent pitch height difference between adjacent lexical tones, not to mention the larger range of possible pitch reset due to individual variations. As such, cases where upwards F_0 adjustment hinders tone recognition can be quite frequent. As for the interaction between pre-boundary lengthening and the contrastive syllabic duration, our analysis indicated that in Cantonese, variations in syllabic duration due to pre-boundary lengthening are simply too small to trigger syllabic miscomprehension. A series of experiments is underway to investigate the perceptual impact of these prosodic interactions. It is our hope that these experiments will shed light on the role of prosody in speech processing in terms of the interrelation between tone and intonation, as well as details about the use of pitch height and syllabic duration as distinguishing features in morpheme recognition.

References

Barrie, M. 2003. "Contrast in Cantonese Vowels", *Toronto Working Papers in Linguistics*, 20: 1-19.

Bauer, R. S. & Benedict, P. K. 1997. *Modern Cantonese Phonology*, Mouton de Gruyter.

Boersma, P. & Weenink, D. 2005. Praat: doing phonetics by computer (Version 4.3.01) [Computer program]. Retrieved from http://www.praat.org/.

Chao, Y.-R. 1930. "A system of tone letters", *La maître phonétique*, 45: 24-27.

Chan, M. K. M. 1999. "Review of: Bauer, Robert S. and Paul K. Benedict (1997). Modern Cantonese Phonology. Berlin and New York: Mouton de Gruyter." *Cahiers de Linguistique Asie Orientale*, 28.1:101-112, The Ohio State University.

Chow, I. 2003. *Prosodic structure in French and Mandarin*, Ph.D. Thesis. University of Toronto.

_____. 2005a. "Resolving temporary syntactic ambiguity by prosodic devices in Mandarin: An Acoustic Study", *Journal of Chinese Linguistics*.

_____. 2005b. "Prosodic cues for syntactically-motivated junctures", paper submitted to the Interspeech 2005 Conference. Lisbon.

Delais-Roussarie, E. 1995. *Pour une approche parallèle de la structure prosodique: Etude de l'organisation prosodique et rythmique de la phrase française*, Ph.D. Thesis, Université de Toulouse-Le Mirail. Toulouse, France, unpublished.

Gandour, J. 1981. "Perceptual dimensions of tone: evidence from Cantonese", *Journal of Chinese Linguistics*, 9: 20-36.

Kwok, H. and Luke, K.K. 1986. 香港粤语语调初探 (An Initial Study of Intonation in Cantonese) 《语文杂志》.

Gussenhoven, C. 1988. "Intonational phrasing and the prosodic hierarchy", Dressler, W. U., Luschutzky, H. C., Pfeiffer, O. E., Rennison, J. R. (eds.), *Phonologica*, 89-99. Cambridge University Press.

_____. & Rietveld, A. C. M. 1992. "Intonation contours, prosodic structure and Preboundary Lengthening", *Journal of Phonetics*, 20(3): 283-303.

Kao, D. L. 1971. *Structure of the Syllable in Cantonese*. The Hague: Mouton.

Ladefoged, P. 1990. "Some reflections on the IPA", *Journal of Phonetics*, 18: 335-346.

Lee, W.-S. 2004. "The effect of intonation on the citation tones in Cantonese", *Proceedings of the TAL2004 Conference*, 119-122.

Li, X.-D. 李行德. 1985. "廣州話元音的音値暨長短對立 [Opposition of vowel quality and length in Cantonese]", *方言[Dialects]*, 1:28-38.

Mai, Y. 麥耘. 1998. "廣州話語調說略 [A brief introduction to Cantonese tones], *廣州話研究與教學 [Research and Pedagogy of Cantonese],* 3:269-280.

Martin, P. 2005. *WinPitch* (Version 1.92) [Computer Program]. Retrieved from http://ww.winpitch.com.

Shen, X. S. 1992. "A Pilot study on the relation between the temporal and syntactic structure in Mandarin", *Journal of the International Phonetic Association*, 22(1/2): 35-43.

Shih, C.-L. 1986. *The prosodic domain of Tone Sandhi in Chinese*, Ph.D. thesis, University of California San Diego.

Swerts, M. G. J. 1997. "Prosodic features at discourse boundaries of different strength", *Journal of the Acoustical Society of America*, 101(1): 514-521.

_____ & Geluykens, R. 1994. "Prosody as a marker of information flow in spoken discourse", *Language and Speech*, 37: 21-42.

Vance, T. 1976. "An Experimental Investigation of Tone and Intonation in Cantonese." *Phonetica*, 33:368-392.

Wightman, C. W., Shattuck-Hufnagel, S., Ostendorf, M., and Price, P. J. 1992. "Segmental durations in the Vicinity of Prosodic Phrase Boundaries", *Journal of the Acoustical Society of America*, 92: 1707-1717.

Xu, Y. 1997. "Contextual Tonal Variations in Mandarin", *Journal of Phonetics,* 25(1), 61-83.

Zee, E. 1993. "A phonetic and phonological analysis of Cantonese vowels and diphthongs", *The Fourth International Conference on Cantonese and Other Yue Dialects,* Hong Kong: City Polytechnic of Hong Kong, 98.

Zeng, X-L., Martin, P. 2004. “Tones and Intonation in Declarative and Interrogative Sentences in Mandarin”, *Paper presented at the International symposium on tonal aspects of languages,* Beijing.

Zhang, N. 1997. “The Avoidance of the Third Tone Sandhi in Mandarin”, *Journal of East Asian Linguistics*, 6(4):293-338.