Prosodic Cues for Syntactically-Motivated Junctures

Ivan Chow

Theoretical and Applied Linguistic Laboratory
University of Western Ontario, London, Ontario, Canada

ichow3@uwo.ca

10.21437/Interspeech.2005-39

Abstract

A pilot study was conducted to examine the manner in which Cantonese speakers use prosody to mark syntactic junctures in speech production. Test sentences were designed to create the intended experimental conditions — each sentence-pair consists of an identical array of morphemes with exactly two interpretations according to two different syntactic structures. These structures can be clarified by conveying prosodic boundaries that coincide with junctures between major syntactic constituents. The purpose of the experiment as well as the meaning of the test sentences (with respect to their syntactic structures) were explained to the subjects. After a few practice runs, they were asked to read these sentences aloud, conveying the marked boundaries, while their voices were being recorded. Three prosodic juncture markers: pauses, pitch reset and pre-boundary lengthening were under examination. Results of the acoustic and statistical analyses of the recorded signals indicated that, pauses were the most effective way of marking junctures, followed by pitch reset. Pre-boundary lengthening was found to be infrequent; in the rare cases where it was detected, pre-boundary syllables were only slightly longer than their non-boundary counterparts. Nonetheless, vast individual differences in terms of the amplitudes and frequencies of prosodic juncture markers were observed. The present study provides acoustical data regarding the manner in which Cantonese speakers use prosody in utterance structure clarification, *vis-à-vis* their specific language experience. In a large-scale experiment, test sentences will be embedded in contextual paragraphs that semantically and prosodically prompt readers to convey the intended prosodic structure. This experiment is underway and is expected to yield more conclusive results.

1. Introduction

Prosody and syntax are separate but related linguistic domains. The relation between them hinges on the observation that "grammatical (syntactic) structure interacts with phonology only to the extent that it determines the prosodic domains delimiting the operation of phonological rules" ([4]:20, cf.[16] etc.). Under this assumption, Chow [5],[6] conducted a series of cross-language acoustic studies to investigate the manner in which French and Mandarin speakers use prosody to mark junctures between major syntactic constituents. The two languages are regarded as phonologically different— one is a non-tonal Indo-European language where melodic movements are defined by tones associated with the last syllable of each prosodic word [14], whereas the other is a lexical tone language in which the melodic contour of an utterance is largely defined by the concatenation of lexical tones associated with each syllable. Despite differences in their respective prosodic structures and the types of linguistic information conveyed by prosody, pitch reset, pauses and pre-boundary lengthening were found to be prominent juncture markers in both languages. Pitch reset refers to the phenomenon where the otherwise continuously declining melodic baseline (e.g. [9], also referred to as 'downdrift') is interrupted at the boundary location. The pitch register after the boundary is readjusted to a higher value where the baseline continues to decline towards the end of the sentence [19]. Pre-boundary lengthening, on the other hand, refers to the phenomenon in which the duration of the syllable immediately preceding the prosodic boundary is increased compared to the average duration of syllables in non pre-boundary locations [10].

As part of a research program designed to investigate the processing of prosodic juncture markers in speech production and perception, a pilot study was conducted to investigate the manner in which Cantonese speakers use prosody to clarify utterance structure. Language-specific differences have been found in terms of the types of information conveyed by prosody, as well as the way speakers use prosodic devices to convey utterance organization (e.g. [5], [22]). Past research in Chinese prosody concentrated on studying the effects of Third Tone Sandhi in Mandarin (e.g. [5], [17], [18], [21], [24]. Although Mandarin and Cantonese are regarded as two closely related languages, Zhan & Cheung ([25]:262) found that only 10% of the 1401 basic words tested were etymologically related and share the same written Chinese characters. In addition, the two languages have very different phonological systems— Cantonese has a larger and more complex lexical tone system that contrasts both with regards to tone shape and relative pitch. It also boasts a syllabic repertoire that contrasts in terms of syllabic duration. Such properties are not present in Mandarin ([1], [23]:98). As a result, a number of interesting questions arise as to whether Cantonese speakers, as a group and as individuals, show any particular behavioural patterns that have not previously been documented, in other words, whether this unique tonal and temporal environment has any effects on the functioning of utterance structure clarification by means of prosody. If so, what are these effects? And what can they tell us about the processing of prosodic juncture markers?

In terms of experimental methodology, although observations made from examining natural conversations are more indicative of speakers' natural behaviour, as the researcher has little or no control over the content and prosodic context of the conversation, the prosodic effects under study are not necessarily elicited. On the other hand, the use of syntactically ambiguous test sentences allows for the creation of experimental conditions that can isolate the

specific effects under scrutiny (cf. [12], [17], [20]). As our experiment aims to isolate the prosodic manifestations of juncture markers from contextual (tonal and temporal) variations created by different Cantonese morphemic combinations, we opted to use a series of specially designed read-aloud test sentences. Specific features of these test sentences are discussed in the section 2. Section 3 presents the experimental methods and data analysis. Test results and their theoretical implications are discussed in section 4, followed by concluding remarks in section 5.

2. Syntactically Ambiguous Test-Sentences

Five pairs of test sentences were designed for this acoustic study. Within each sentence pair, corresponding sentences consist of an identical array of morphemes. In Cantonese, a morpheme consists of a syllable combined with a lexical tone. Most morphemes on their own are treated as words and have their own lexical definitions, while others obtain their meanings only when combined with one or two other morphemes. Within the test sentences, the boundary between major syntactic constituents (e.g. the subject NP and the VP predicate) can be placed at two different locations. Thus the meaning of the sentence is changed when the boundary location is altered. In each sentence pair, syntactic boundaries are placed either between syllables 2 and 3 (as well as between 4 and 5 in one case), or between syllables 3 and 4 (see figure below). This way, prosodic changes can be analyzed by direct comparison—where a boundary is present after syllable 2 in structure (a), it is not present at the same location in structure (b). Melodic and temporal variations at this particular location can then be measured and compared directly against the corresponding sentence to determine the way in which the syntactic boundary was prosodically conveyed. While using identical series of morphemes in both alternatives, contextual variations (in terms of pitch and syllabic duration) in different morpheme combinations were kept to a minimum. As a result, measured acoustic correlates can be readily attributed to prosodic juncture marking.

Figure 1 shows the syntactic structures of a sample sentence pair. (The P.T. Lau system is used for Cantonese segmental transcription; a revised version of the Y.R. Chao Tone Numbers is used for tonal transcription. Refer to [1] for more discussions on Cantonese romanization and tone numbers.)

(a) Jeung55 man21 / sue55 se25 man21 jeung55.
張文/書寫文章。
"Zhang-wen writes articles."

(b) Jeung55 man21 sue55 / se25 man21 jeung55.
張文書/寫文章。
"Secretary Zhang writes articles."

(a) (b)

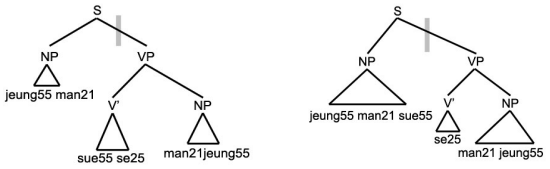


Figure 1. Sample of a pair of test sentences and their respective syntactic structures

In figure 1, the opaque line denotes boundaries between two major syntactic constituents. This sentence was designed in such a way that the third syllable can be part of the VP-predicate as in structure (a); or it can be part of the NP-subject as in structure (b). Depending on the (syntactically-motivated) prosodic grouping, the sentence has exactly two interpretations. No other permutations of boundary location would render sensible interpretations other than the ones listed above.

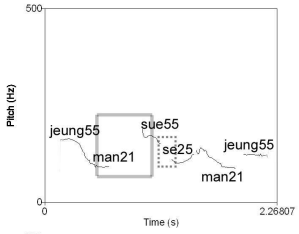
3. Experimental Methods

Three subjects volunteered to take part in the experiment. The subject pool consisted of two male and one female, aged 31-65, all native speakers of the Hong Kong variety of Cantonese, with no known hearing and speech impediments. Subjects were asked to sit in a quite room and were given a printed script containing the test sentences. The semantic differences of the test sentences and the purpose of the marked boundaries were explained to them. After a few practice runs, they were asked to read the script aloud as if they were engaged in a natural conversation, while their voices were being recorded using a digital sound recorder. Acoustic analyses were conducted on the recorded signals using WinPitch [15] and PRAAT [2]. The purpose of the acoustic analyses was to determine the presence or absence of the three prosodic juncture markers in each boundary location, as well as to measure their amplitudes.

3.1. Pitch reset

Pitch reset refers to the interruption of the otherwise continuously declining melodic baseline; the pitch register of the declination line is adjusted to a higher point where it continues to decline towards the end of the sentence.

(a).



(b).

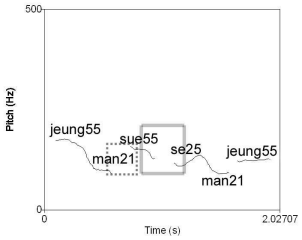


Figure 2. Measurement of Pitch Reset

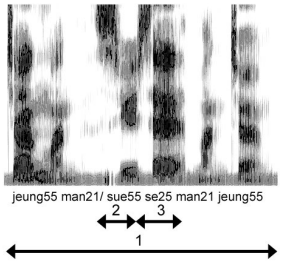
In figure 2, the (larger) solid boxes denote the locations of syntactic boundaries and the segments of pitch contours under scrutiny, whereas the (smaller) dotted boxes represent the pitch contour segments at the same location within the corresponding 'non-boundary' sentences. Following Swerts [19], the amplitude of pitch reset is measured by way of pitch range difference (PRD). Pitch range (measured in Hz) is measured at the nuclear vowel segment of the syllables before and after the boundary at the maximal point of vowel intensity. PRD is derived by subtracting the pitch range of the post-boundary syllable from that of the pre-boundary syllable.

Since, in addition to pitch reset, PRD can be influenced by contextual variations (cf. [21]) of lexical tones in Cantonese, measured PRD values were compared between the two alternative sentences at identical locations. When the PRD of a boundary location (solid boxes) was greater than that of its non-boundary counterpart (dotted boxes), pitch reset was determined to be present as a juncture marker.

3.2. Pre-boundary Lengthening and Pauses

Pre-boundary lengthening has been reported as a prosodic juncture marker in different languages ([10] for Dutch; [5] for French; [6] and [17] for Mandarin).

(a).



(b).

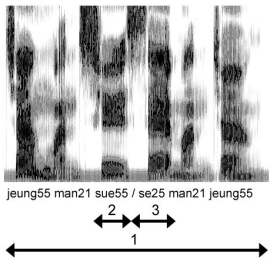


Figure 3. Measurement of Pitch Reset

In this study, syllabic duration of pre-boundary syllables in both (a) and (b), i.e. syllables 2 and 3, were measured in milliseconds (denoted 2 and 3 in the figure). These measurements were then divided by the duration of the entire sentence (1) to compensate for speech rate difference between corresponding sentences. Subsequently, the percentage values were compared to determine whether pre-boundary lengthening was present at a boundary location. Since syllabic duration is a contrastive feature [11], [13] in Cantonese, the comparative method used here helped to minimize the impact of contextual variations in different morpheme combinations that would otherwise be present had recorded natural conversations been used.

Pauses were measured and compared between the two corresponding sentences in the same fashion. Some small inter-syllabic gaps were found at the same location in the corresponding 'non-boundary' sentences, durations of these gaps were nonetheless duly noted and used in the comparative analysis.

4. Results and Discussions

A total of 33 boundaries (11 boundaries within the test sentences x 3 subjects) were examined. Figure 4 shows the overall frequencies of each prosodic juncture marker.

Prosodic Juncture Markers	Frequency
Pitch Reset	88%
Pre-boundary Lengthening	64%
Pauses	97%

Figure 4. Overall Frequencies of the Prosodic Devices

Among the three prosodic devices under investigation, pauses were the most prominent juncture marker, appearing in all but one instance of syntactic boundaries. The durations of pauses were on average 17.9% (of the overall sentence duration) longer than the gaps (if any) between the same syllables in the corresponding sentences, which in most cases measured close to 0ms. Pitch reset was the second most prominent prosodic marker. It was detected in 88% of the boundaries under investigation. When pitch reset was detected, average pitch range difference across boundaries measured 15.4Hz (SD=12.5Hz). In other words, at boundary locations where pitch reset was present, the pitch register of the post-boundary syllable was on average 15.4Hz higher than the same syllable had there not been a boundary. Pre-boundary lengthening, on the other hand, was shown to be less frequent than the other two prosodic devices. It was detected in only 64% of the cases. Moreover, durations of the pre-boundary syllables were on average only about 1.8% of the overall sentence duration longer than their non-boundary counterparts. While individual subject analyses showed behaviour in pauses and pitch reset to be consistent between all three subjects, vast differences were found in terms of the use of pre-boundary lengthening.

Subject	Freq (%)	% dif. in syllabic duration (%)	S.D. (%)	Range (%)
IHM	91%	1.3%	0.7%	0.4% -2.8%
THF	73%	2.7%	2.8%	0.5% -9.0%
SHM	27%	1.5%	2.0%	0.3% -4.0%

Figure 5. Individual Data for Pre-boundary Lengthening

Although pre-boundary lengthening was frequently used by subjects IHM and THF, it was used by subject SHM only 27% of the time. If data from subject SHM were taken out of the analysis, the frequency of pre-boundary lengthening jumps to 82%, much closer to those of the other prosodic markers; lengthened syllables were on average 2% of the overall sentence duration longer than their non-lengthened counterparts. However, overall range also revealed that in some cases the pre-boundary syllables could be lengthened by up to 9% of the overall sentence duration.

All in all, results from our data analysis showed pauses to be by far the most prominent prosodic marker of syntactic junctures. This result is consistent with the findings of a similar study conducted by [17] with Mandarin speakers. As Chow [5], [6] showed pitch reset to be a prominent juncture marker in Mandarin and French, the present study yielded consistent results for Cantonese speakers— to the author's knowledge, it is the first documentation of pitch reset as a melodic juncture marker in Cantonese.

As for pre-boundary lengthening, our results showed that not only was it significantly less frequent as a juncture marker, in the rare cases where its was detected, the duration of the lengthened syllable was on average only slightly longer than its non-boundary counterpart. Yet, two out of three subjects used pre-boundary lengthening rather frequently, and the pre-boundary syllables could be lengthened quite

significantly. It would be of interest to determine whether the percentage of increase in syllabic duration derived in this study has any perceptible effects as a juncture cue in the parsing of utterance structure.

5. Conclusion

A pilot study was conducted to examine the manner in which Cantonese speakers use prosody to mark syntactic junctures in speech production. Acoustic analysis of the recorded test sentences revealed that, pauses and pitch reset were significantly more frequent and prominent than pre-boundary lengthening.

Individual variations played an important role in the statistic analysis of experimental data. Although results showed consistent behavioural patterns as those reported in studies on Mandarin and French [6] in terms of pauses and pitch reset, subjects behaved quite differently in marking prosodic junctures using pre-boundary lengthening.

Although an anonymous reviewer indicated that using an explicitly marked script may create a testing environment vastly different from that of a natural conversation, it is not uncommon in laboratory phonology to use specially designed scripts to elicit the prosodic/phonological effects under study. Especially when these effects cannot otherwise be studied using recorded conversations. Nonetheless, in order to improve the naturalness of the elicited response, the test sentences used in a large-scale study will be embedded within contextual paragraphs that will semantically and prosodically prompt readers to convey the intended structures. This study is underway and is expected to yield more indicative results towards the understanding of the prosodic cues used by speakers of different languages to process utterance structure.

Given the fact that Cantonese speakers use pitch and rhythm to mark syntactic junctures, an interesting question arises as to whether melodic and temporal variations created by juncture marking would interfere with the comprehension of Cantonese morphemes.

6. Acknowledgements

I am indebt to Marc Swerts and Carlos Gussenhoven for their invaluable suggestions on the development of this research program. I would also like to thank the anonymous IS2005 reviewers for their insightful comments on this paper.

7. References

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

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

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

[4] Chen, M. Y. "What Must Phonology Know about Syntax?" Sharon Inkelas and Draga Zec (eds.), *The phonology-syntax connection*, 19-46, Chicago: University of Chicago Press, 1990.

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

[6] _____. "Resolving temporary syntactic ambiguity by prosodic devices in Mandarin: An Acoustic Study", *Journal of Chinese Linguistics*, 2005 no. 33, vol. 2.

[7] _____. "Interactions between Cantonese Syllabic-Level Prosody and Prosodic Juncture Markers within the IP". Unpublished Manuscript.

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

[9] Gussenhoven, C. "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, 1988.

[10] _____ & Rietveld, A. C. M. "Intonation contours, prosodic structure and preboundary lengthening", *Journal of Phonetics*, 20(3): 283-303, 1992.

[11] Kao, D. L. *Structure of the Syllable in Cantonese*. Mouton, the Hague, 1971.

[12] Lehiste, I. "Phonetic disambiguation of syntactic ambiguity", *Glossa*, 7: 107-122, 1973.

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

[14] Martin, P. "L'intonation du français et du portugais: phonétique et phonologie", *Revista de l'AMPOLL*, Sao Paulo, 5/7:255-295, 1999.

[15] _____. *WinPitch* (Version 1.92) [Computer Program]. Retrieved from http://www.winpitch.com, 2005.

[16] Selkirk, E. O. *Phonology and Syntax: the Relation between Sound and Structure.* Cambridge, MA. MIT Press, 1984.

[17] Shen, X. S. "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, 1992.

[18] Shih, C-L. "Mandarin Third Tone Sandhi and Prosodic Structure", J. Wang and N. Smith (eds). *Studies in Chinese Phonology,* Mouton de Gruyter. 81-123, 1997.

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

[20] Warren, P. "Prosody and Parsing: An Introdution", *Prosody and Parsing*, 1-16, Hillsdale, NJ & Hove: Lawrence Erlbaum Associates. 1996.

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

[22] Yaegor-Dror, M. "Register and Prosodic Variation, a Cross-Language Comparison", *Journal of Pragmatics: An Interdisciplinary Journal of Language Studies*, 34, no. 10-11: 1495-1536, 2002.

[23] Zee, E. "A Phonetic and Phonological Analysis of Cantonese Vowels and Diphthongs", *Abstracts, the 4th International Conference on Cantonese and other Yue Dialects*. Hong Kong: Linguistic Society of Hong Kong: 98, 1993.

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

[25] Zhan, B.-H. 詹伯慧 & Cheung, Y.-S. 張日昇. "珠江三角洲粵方言的常用詞 [Pearl River Delta Dialects: similarities to and differences from the Beijing dialect]", *方言[Dialects*], 4: 262-266, 1989.