

Accentual Lengthening of Monosyllabic-Constituents in Beijing Mandarin

Yiya Chen

Linguistics Department
State University of New York at Stony Brook
Yiyachen@yahoo.com

Abstract

In stress-accent languages such as English and Dutch, pitch accent is used as the primary cue for contrastive focus, a mechanism to indicate correction or emphasis in discourse. A well-known fact about Chinese is that pitch change is used to indicate tones for lexical contrasts. This study examines how an alternative acoustic cue - duration - is used to signal contrastive focus in Beijing Mandarin. In particular, attention is paid to the interaction of accentual lengthening with other known factors that govern segment duration, namely vowel intrinsic duration and a segment's position in the prosodic hierarchy. Results from one-syllable constituents will be presented to show their conjoint influence.

1. Introduction

Studies on focus in stress-accent languages such as English and Dutch found that while an increase in both duration and intensity are observed to accompany focus, F0 is the most reliable acoustic realization of focus (see [1] and references therein). This raises the question of how focus is conveyed in Beijing Mandarin, a tone language, where F0 is used to convey lexical contrasts. This study therefore examines another acoustic cue - duration - and reports experimental results on the function of this particular cue in signaling focus. Of particular interest here is that segment duration is also governed by various other factors such as vowel, stress, tone, and sentential position (see [2] and [3] for an overview). We will examine the basic patterns of accentual lengthening observed on monosyllabic constituents in Beijing Mandarin. Three specific questions are asked: 1) when a monosyllabic constituent is focused, which part of the syllable lengthens? 2) what is the interaction of accentual lengthening with vowel intrinsic duration? Specifically, is vowel intrinsic duration maintained under contrastive focus? 3) what is the magnitude of accentual lengthening as a function of sentential position of the syllable under focus?

2. Methods

The general purpose of this experiment was to see what acoustic cues are employed by Beijing Mandarin speakers to convey contrastive focus and how these

cues are implemented. The procedures used in this experiment involved the oral reading of different sentences in which the location of contrastive focus was systematically varied.

2.1. Data design

The basic test sentences used in this study are exemplified in (1).

- (1) **X** **Y**
 | |
tā shuō shuō mā nán duō le.
He say say mother difficult more Aspect
'He said to say mother is more difficult.'

The frame sentence used in this study is *tā shuō shuō X Y duō le* (the variables are indicated by the bolded letters). *X* is replaced by eight syllables: *ma* and *mi* with four lexical tones in Mandarin respectively. *Y*, representing the syllable that follows *X*, also has two versions, one with a rising tone *nán* 'difficult' and the other with a falling tone *màn*, 'slow'. For each sentence, there are four focus conditions, which are illustrated in (2). Underlined syllables with subscripts are the ones that are elicited with focus at different times. There is only one focused item per utterance.

- (2) Four focus positions:
e.g. tā shuō shuō₁ mā₂ nán₃ duō₄ le.
He say say mother difficult more Aspect
'He said it is more difficult to say mother.'

2.2. Elicitation of focus

Focus was elicited by providing subjects with relevant contextual information. Subjects saw the sentence (3) on the computer screen. They were also shown the question in (4). A typical answer from the subject is shown in (5). All sentences were given in Chinese.

- (3) Sentence
a. tā shuō shuō mā nán duō le.
he say speak mother difficult more asp.
'He said that it is more difficult to say mother.'
b. tā shuō xiě mā nán duō le.
he say write mother difficult more asp.
'He said that it is more difficult to write mother.'

(4) Question: Suppose he said sentence (a), and I said he said sentence (b), how would you utter (a) to correct me?

(5) Response: tā shuō shuō mā nán duō le.

2.3. Subjects and recording procedures

Data reported here are from 1 male and 2 female native speakers of Beijing Mandarin. All were born and grew up in Beijing. All are graduate students and have been in the United States for less than 5 years. The sentences were blocked by focus condition. The two female subjects were asked to repeat the task five times, each time with the same order for the four blocks and random order within the block. This turned out to be a physically quite challenging task and the male speaker produced only four repetitions. Recording was done with Sony Digital Mega Bass MZ-R55 at the sampling rate of 44100. All recording sessions were carried out in the sound booth of the Phonetics Lab at the State University of New York at Stony Brook. All subjects were told that the purpose of the recording was a study of focus in Mandarin, but they were naïve as to what exactly would be examined. During the recording, whenever subjects responded with focus on the wrong word, they were asked by the experimenter to repeat the utterance. However, if they did not emphasize enough, the recording would just go on for the purpose of not interrupting their overall pattern of production.

2.4. Acoustical analysis

The recordings were down-sampled to 16,000Hz in GoldWave. They were analyzed in CSL. Segmentation was based on both spectrogram and waveforms, with reference to the standard criteria of segmentation. Duration was examined as a possible cue for focus. Although there are four foci in the data collected, only the three words shown in (6) were measured for duration. *Duo* was excluded simply because in most cases, there is no clear boundary between *duo* and the following word *le*. Among the three words, for reasons to become explicit in the data analysis session, ten sets of durational measurements were taken from the segmented label files: duration of the three words (as illustrated in (6) where *W* represents word), as well as duration of the onsets, rhymes, and codas (applicable only for the third word) of these three words.

(6) W1 W2 W3
 | | |
 tā shuō shuō mā (X) nán (Y) duō le.
 He say say mother difficult more Aspect
 'He said to say mother is more difficult.'

2.5. Statistical analysis

Statistical analysis was performed on the coded data using SPSS to determine the main effects of contrastive focus on duration. Analysis of variance was performed on all or some of the following factors for the three words of interest: tone, vowel, following word, speaker, and focus type. Separate ANOVAs were performed on the three words of interest because the segmental, tonal make-up of these words as well as

their position in the sentence were very different. Here, I will show the individual results only when relevant to the following discussion.

3. Results

3.1. Word as a unit of accentual lengthening

Statistical results show that all three words lengthened significantly under contrastive focus (Table 1).

Table 1: Results of ANOVA for the effect of accentual lengthening.

Shuo	Onset	$f(3, 460) = 278.622$	$p < .001$
	Vowel	$f(3, 460) = 161.463$	$p < .001$
Ma	Onset	$f(3, 460) = 75.049$	$p < .001$
	Vowel	$f(3, 460) = 77.318$	$p < .001$
Nan	Onset	$f(3, 460) = 23.083$	$p < .001$
	Vowel	$f(3, 460) = 88.286$	$p < .001$
	Coda	$f(3, 460) = 63.616$	$p < .001$

Focus affects the duration of all subcomponents of a syllable. A post-hoc scheffe test suggested that when a word is under focus, it is significantly longer than under the other three conditions (preceding, following, or nonadjacent to focus). However, among the three non-focus conditions, duration doesn't differ much and no consistent significant results were found. Subject was also found to be a significant factor in accentual lengthening, with individual variations in terms of the accentual lengthening pattern (figures 2-4).

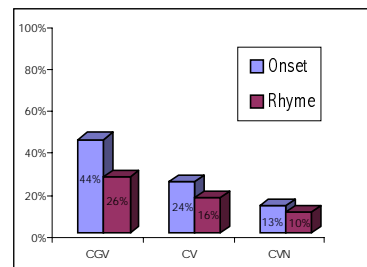


Figure 2 Speaker One: More onset lengthening

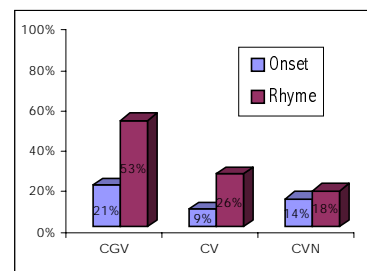


Figure 3 Speaker Two: More vowel lengthening

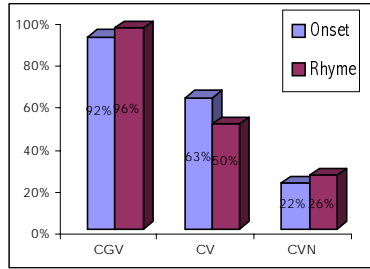


Figure 4 Speaker Three: No consistent pattern

Figure 2 shows that Speaker One consistently lengthens the onset more than the rhyme. Figure 3 shows that Speaker Two consistently lengthens the rhyme more than the onset. Speaker Three (figure 4) is less consistent. In general, he lengthens much more than the other two speakers under contrastive focus but lengthens the onset more in some cases and the rhyme in others. These individual variations seem to suggest that speakers do have at their disposal different strategies to organize temporal structures under contrastive focus.

3.2. Intrinsic vowel duration under accentuation

The result of ANOVA with vowel, tone, and focus as fixed factors show that there is no significant interaction between vowel quality and focus. I thus interpret this as support for the claim that intrinsic vowel duration differences are maintained under contrastive focus. Worth noting is that these length differences are only significant for some tones. The pattern, however, is consistent across speakers

Ancillary to the examination of intrinsic vowel duration differences is the finding that onset and vowel exhibit a complementary lengthening pattern (c.f. [3]), i.e. while the low vowel *a* is consistently longer than the high vowel *i* across different tones and focus conditions, the onset of *ma* is consistently shorter than the onset of *mi*.

3.3. Magnitude of lengthening as a function of sentential position

Figures 2-4 exhibits another aspect of the lengthening pattern: the magnitude of accentual lengthening appears to decline over a sentence. The earlier a syllable is in a sentence, the more it is lengthened under contrastive focus.

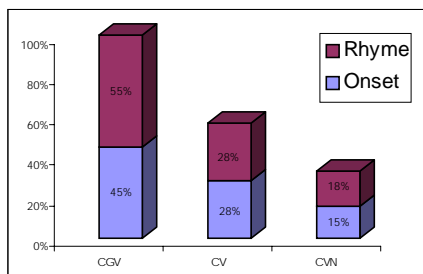


Figure 5 Magnitude of lengthening: pooled across speakers

This declining magnitude of accentual lengthening may be questioned as the syllables here are not identical and they differ in their internal structures. However, possible supporting evidence, is found in two other studies. Neither study looked into the pattern of accentual lengthening, but they did report the lengthening effect induced by sentence stress.

Xu 1999 ([4]) studied the peak alignment of tone in Mandarin. One example of the sentences that he used for the study is shown here. Note that in this study, he consistently used bisyllabic monomorphemic words in both the first and the last positions and a monosyllabic word in the middle. Although the syllable structures and segments are not always the same, they are very similar. I took the duration values reported in his paper and recomputed the percentage of lengthening (figure 5). Here we again see a declining pattern.

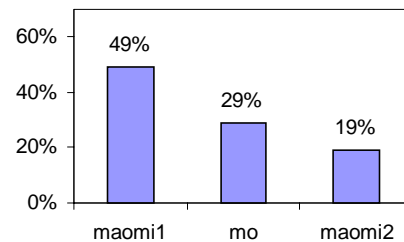


Figure 6 Accentual lengthening from Xu (1999)

Another study is Jin 1995 ([5]) which examined the different acoustic realization of broad vs. narrow sentence stress. I again, recomputed the percentage of lengthening with the durational data reported there. As chart 13 shows there is also declination trend.

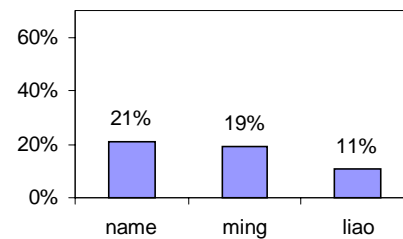


Figure 7 Accentual lengthening from Jin (1995)

4. Discussion

The results of this experiment show that speakers of Beijing Mandarin do employ duration, as an acoustic cue, to convey contrastive focus. Furthermore, the data show that 1) both onset and rhyme contribute to the durational changes caused by accentuation and 2) speakers exhibit robust individual variations. There are two possible lines for interpretation. First, this could be taken as evidence that the whole syllable as a word is a relevant unit of tempo in Beijing Mandarin ([6] and [7]). The complementary durational effects of onset and vowel found in this study lend further

support to the proposal that it is not a subcomponent of a syllable, but the whole syllable which is the relevant unit of tempo in Beijing Mandarin¹. This makes it interesting to see how speakers organize their temporal structures when a bi-syllabic constituent is under contrastive focus. If word indeed is a relevant unit of tempo, we would then predict different patterns of accentual lengthening on bi-syllabic constituents varying according to their morphological status as a word or a phrase composed of two independent words. Second, it could also be that the accentual lengthening exhibited here should be attributed to two sources: one is the lengthening of the rhyme due to accent; and the other is the lengthening of onset due to the initial strengthening effect, resulting from a new domain created by focus ([8]), which is higher on the prosodic hierarchy than the domain where the word was without focus. This, however, also awaits further study of accentual lengthening on constituents larger than one syllable.

In this study, we also observed a declining magnitude of accentual lengthening. Admittedly, none of the studies mentioned above were designed to see the sentential effect on the magnitude of accentual lengthening. However, we do see a declination pattern emerging. There are two accounts for declination. The traditional explanation is that declination is a global weakening effect of serial position in a sentence and is caused by a weakening of glottal or supralaryngeal articulation. A more recent proposal is made by Fougeron and Keating 1997 ([8]). They studied the articulatory strengthening at edges of prosodic domains and proposed that declination could be a more local effect, resulting from serial position within any given prosodic domain, rather than a sentence. Given the alternative account for declination, it remains to be seen why the lengthening effects of contrastive focus are different as a function of the syllable's serial position in a sentence. Prosodically speaking, the three syllables examined here, despite their locations within a sentence, should be at the edge of the same prosodic domain ([10]). Several questions arise for further research, such as 1) why does the degree of lengthening vary with position? 2) what role do these differences play in perception of focus?

5. References

- [1] Sluijter, Agatha, 1995. *Phonetic Correlates of Stress and Accent*. Dordrecht.
- [2] Van Saten. J.P.H., 1992. Contextual effects on vowel duration. *Speech Communication* 11(6), 513-46.
- [3] Shih, Chilin; Benjamin Ao, 1997. Duration study for the Bell laboratories Mandarin Text-to-Speech System. In *Progress in speech communication*. Van Saten, Jan. P.H., Richard W. Sproat, Joseph P. Olive & Julia Hirschberg (eds.). NY: Springer.
- [4] Xu, Y., 1999. Effects of Tone and Focus on the Formation and Alignment of F0 Contours. *Journal of Phonetics* 27(1): 55-105.
- [5] Jin, Shunde, 1996. An acoustic study of sentence stress in Mandarin Chinese. Doctoral dissertation, Ohio State University.
- [6] Eafiting, W., 1991. The effect of information value and accentuation on the duration of Dutch words, syllables, and segments. *JASA* 89. 412-24.
- [7] Cambier-Langeveld, Tina, 2000. Temporal marking of accents and boundaries. Netherlands Graduate School of Linguistics.
- [8] Shih, Chilin, 1989. Mandarin third tone sandhi and prosodic structure. In Wang Jialing and Norval Smith, (eds.) *Studies in Chinese Phonology*. Foris, Dordrecht.
- [9] Fougeron, C.; P. A. Keating, 1997. Articulatory strengthening at edges of prosodic domains. *JASA*. 106(6). 3728-40.
- [10] Beckman, Mary; Pierrehumbert, Janet, 1986. Intonational structure in English and Japanese. *Phonology Yearbook* 3: 255-309.

Acknowledgement

This work was supported in part by NSF grant SBR 9600930 to Marie Huffman. Thanks to Ellen Broselow and Marie Huffman for questions and guidance, and to Chihlin Shih and Yi Xu for discussions. Without my subjects, this work would have been impossible.

ⁱ As one of the reviewers pointed out, the strong individual variation between the speakers also deserves attention, either to reject or reinforce with a greater sample in future research.