

AN ACOUSTIC STUDY OF THE INTERACTION BETWEEN STRESSED AND UNSTRESSED SYLLABLES IN SPOKEN MANDARIN

Jing Wang

Speech, Hearing and Language Research Centre,
Macquarie University, Sydney

ABSTRACT

This study examines the acoustic correlates of relative prominence in connected Mandarin speech, particularly those acoustic manifestations that reveal an interaction between stressed and unstressed syllables. The acoustic measurements are primarily concerned with segment duration, intensity, and F0. This study shows that: (1) stressed CV syllables have significantly longer consonant and vowel durations and greater intensity when preceding unstressed syllables than stressed syllables; (2) vowel intensity and duration of stressed syllables are significantly affected by the stressed status of preceding and succeeding syllables, and anticipatory effects are greater on vowel duration but carryover ones are greater on vowel intensity; (3) apart from their lexical stress status, the relative prominence of consecutive syllables is mainly determined by interaction—i.e., unstressed syllables render neighboring syllables more prominent, while stressed syllables cause neighboring syllables to become weak, engendering a pattern of alternating strong and weak syllables. This study also provides insights into the metrical structure of Mandarin and how stress functions in Mandarin speech, where stress and accent interact with lexical tones.

1. INTRODUCTION

Stress in a tonal language is sometimes difficult to define. In Mandarin, the existence of lexical stress is a contentious issue that has been the focus of on-going intellectual debate, not only in terms of defining word stress and its linguistic functions, but also its phonological classification and phonetic realization [1, 4, 5, 6, 7, 8, 11]. In contrast to the problematic status of stressed syllables, lexically unstressed syllables in spoken Mandarin are well-formed, and are signaled by a shortening in vowel duration, vowel reduction and the loss of their original lexical tones.

Whether lexical stress exists in Mandarin remains a contentious issue and little research has been conducted on stress and the degrees of stress in Mandarin, either in terms of general discussion on [1, 4, 5, 7] or the acoustic investigation of stress [6, 8, 11]. Some fundamental questions remain unsolved. Firstly, how do we define lexical stress in Mandarin? Can it be seen as the sole of the domain of unstressed, neutral tone syllables? If that were the case, then Chao's [1] three level hierarchy would be the logical starting point, these being: (1) weak stress (unstressed); (2) normal stress; and (3) contrasting stress. However, auditory perceptions indicate that syllables belonging to (2) of Chao's schema are not equally stressed, and therefore require further subclassification for them to be considered accurate. Hence, a system using three degrees of lexical stress has commonly been adopted,

despite the fact that different terms have been used to label the different degrees of stress—i.e., unstressed, weak-stressed and strong-stressed syllables. Furthermore, it has been claimed that the difference in prominence that occurs with normal stressed syllables is due to their positioning, with fringe syllables being more prominent than interior syllables and final syllables being more prominent than initial syllables [1]. This claim seems to be supported by the acoustic results in several studies which show that final syllables have longer duration, a wider pitch range, and a lower bottom pitch [8, 11]. However, there are doubts as to whether acoustic manifestations accurately reflect differences in prominence, and not intonational boundary effects, as most of the research has been conducted using isolated polysyllabic constituents [11]. Acoustic studies using continuous speech have been carried out by Kratochvil [4, 5, 6]. The results of these studies suggest that stress patterns of polysyllabic constituents are relatively unstable, and their relevant prominence is partly determined by a tendency to form alternating strong and weak syllables. His research provides a very different insight into Mandarin stress and rhythm, both phonologically and acoustically.

Motivated by a desire to further understand Mandarin stress patterns in connected speech, which would provide insights into a metrical approach for Mandarin rhythm, this study focuses on the acoustic correlates of relative prominence of syllables in polysyllabic prosodic constituents. Of particular interest is the interaction between consecutive syllables, in terms of their different stressed status, in order to know whether the relative prominence of a syllable is affected by neighboring syllables. If the prominence of Mandarin syllables is relatively unstable, whether due to contexts, or a tendency to form alternating strong and weak syllables, then Mandarin rhythm, in a sense, could be seen as being partly stress-based.

2. MATERIALS AND PROCEDURES

24 sentences containing two polysyllabic (either disyllabic or trisyllabic) constituents with different stress patterns, were read by four native Beijing Mandarin speakers a total of six times. The stress patterns were represented by two degrees of phonological stress: normal stressed (S) and atonic unstressed (U). Sentences were written in Chinese characters on 3" x 5" cards arranged randomly. Subjects, including the author, were asked to read sentences at their natural speech rate. Once the 24 sentences were read, the subjects were asked to repeat the task a total of six times. The speech samples were recorded in the studio at the Speech, Hearing and Language Research Centre at Macquarie University using DATs. Speech samples were then digitized on a SUN workstation at a sample rate of 20 kHz. Acoustic-phonetic segmentation and labeling were manually carried out with Waves+

program and related acoustic data were tracked with mu+ system for speech corpus analysis [3] The following table provides the polysyllabic words used in the sample sentences. An example of this is *qi1fu lao3nian2ren2*, which contains the verb *qi1fu* and noun *lao3nian2ren2*. In the V/N series, verbs are either *qi1fu* or *qi1zha4*, while in the N/V series, verbs are either *chille* or *chi1kui1*. The 24 sample sentences include each combination.

SU (verb)	SS (verb)	SU (noun)	SS (noun)
<i>qi1fu/chille</i>	<i>qi1zha4/chi1kui1</i>	<i>lao3po</i>	<i>lao3ren2</i>
SSS (noun)	SSU (noun)	SUS (noun)	SUU (noun)
<i>lao3nian2ren2</i>	<i>lao3shiren3</i>	<i>lao3ma1ma</i>	<i>lao3shide</i>

3. EFFECTS OF SUCCEEDING SYLLABLES

The effects of succeeding syllables, in terms of their stressed status, were examined by comparing segment duration, F0, intensity, and the spectral quality of vowels contained in the syllables *qi* and *chi*, in both SU and SS contexts.

3.1. Vowel Duration

For all subjects, vowel duration in both *qi* and *chi* is significantly longer when preceding unstressed syllables than stressed ones. The lengthening degree is about 58.1% on average, but is greater in the *chi* series than in the *qi* series, 82.8% and 33.6% respectively. For example, for subject LG, the vowel in *qi* is much longer when preceding unstressed syllables, with a mean duration of 100 ms and a standard deviation of 16 ms, than preceding stressed syllables, with a mean duration of 65 ms and 12 ms SD ($t(70)=19.67, p<0.00001$).

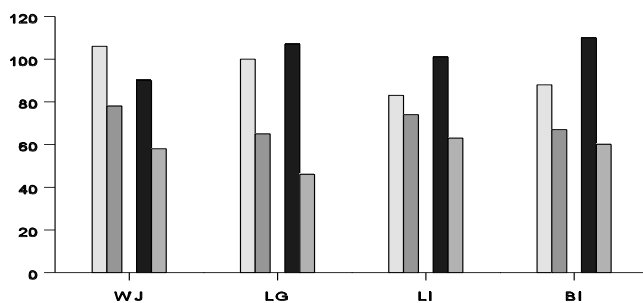


Figure 1: Mean vowel duration in a), *qi* (SU); b), *qi* (SS); c), *chi* (SU) and d), *chi* (SS) for each of the four subjects.

3.2. Vowel Intensity

The overall intensities of vowels are generally greater when preceding unstressed syllables than stressed ones. This phenomenon is very evident in the *qi* series, where target syllables are at the beginning of sentences, but is less apparent in the *chi* series, in mid-sentence. For a quantitative comparison, the intensities at the middle point of vowels are calculated (in dB). The following table provides mean intensity values and standard deviations for each of the four subjects.

	N	<i>qifu</i> (SU)		<i>qizha</i> (SS)		<i>chile</i> (SU)		<i>chikui</i> (SS)	
		M	SD	M	SD	M	SD	M	SD
WJ	36	69.2	3.1	65.2	1.6	69.0	4.0	67.0	3.7
LG	36	72.6	1.8	69.9	1.3	70.0	1.8	68.6	5.2
BI	36	75.9	1.5	72.4	1.4	75.8	2.4	74.9	1.8
LI	36	71.6	1.9	67.4	1.9	69.4	2.4	71.5	2.1

3.3. Fundamental Frequency (F0)

The first tone (high level) syllables, both *qi* and *chi*, were chosen to reduce the effect of different F0 being associated with lexical tones. However, tone coarticulation is still possible—hence only F0 values at middle points of vowels are compared. The results show that the differences in F0 are not consistent in the *qi* series but are fairly consistent in the *chi* series, where F0 tends to be slightly higher when vowels precede unstressed syllables than stressed ones (7 Hz on average, but this is not consistent across the four subjects).

3.4. Spectral Quality

A spectral analysis was conducted to find out if there was any difference in vowel quality in terms of F1/F2 space, when vowels are followed by syllables of different stress status. Not only did single F1 and F2 values fail to consistently show any significant differences, but the $\log(F2/F1)$ measurement, which is supposed to reflect a kind of vowel undershoot or centralization [2], also failed to consistently demonstrate any significant difference.

3.5. Discussion I

The results show that vowels in normal stressed syllables tend to have much longer durations, and relatively greater intensities when they precede unstressed syllables, rather than stressed ones. If the difference was confined to duration, we could interpret it with the compensatory mechanism—i.e., vowel duration increases when succeeding syllables are unstressed, rather than stressed, in order to keep the duration of polysyllabic constituents relatively constant, although the complete compensation is not achieved (SU is 14% shorter, on average, than SS constituents in this study). We found with Mandarin, however, that significant duration lengthening is associated with a significant increase in intensity. Moreover, the two parameters seem to compensate each other to a certain extent. For example, there is a very significant difference in vowel duration (66.3% on average) but little difference in intensity (0.6 dB) in the *chi* series, while a significant but smaller difference in duration (26.5%) but a very significant difference in intensity (3.6 dB), occurred in the *qi* series. All this suggests that vowels are more prominent when they are followed by unstressed syllables than they are when followed by stressed ones, in terms of vowel duration and/or intensity.

4. EFFECTS OF PRECEDING SYLLABLES

Since polysyllabic lexical words and prosodic constituents in Mandarin rarely begin with unstressed, neutral tone syllables, the effect preceding syllables have in terms of their stress status always stems from the final syllables of the preceding prosodic constituents. In the following

section, we examine this effect on the syllable *lao* in V/N sentences.

4.1. Consonant and Vowel Duration

The duration of consonant *l* is slightly longer when it follows an unstressed syllable than it is when it follows a stressed one. However, the significance of this difference varies from subject to subject. For subject WJ, there is a slight significant difference —10 ms (65 ms vs 55 ms; $t(69)=6.83$, $p<0.0001$), while for other three subjects, the difference is not significant at all ($p>0.05$). Further, the inconsistent differences in the duration of the vowel *ao* indicates that the prominence of preceding syllables does not exert much influence on the durations of ensuing segments.

4.2. Consonant and Vowel Intensity

Unlike duration, the effect on the intensity of subsequent segments is both apparent and significant. The overall intensities of consonants and vowels demonstrate that voiced consonants have a much greater intensity when they follow an unstressed syllable than they do when they follow a stressed one. The same applies to vowels, although any increase in the intensity of vowels is mainly located in the first half of those vowels. For a quantitative comparison, the average intensity values at the middle of consonants and the onset of vowels (10%) are given in the following table.

		Consonant <i>l</i>				Vowel <i>ao</i>			
		SU -		SS -		SU -		SS -	
	N	M	SD	M	SD	M	SD	M	SD
WJ	36	67.9	2.8	62.8	2.1	72.0	2.1	65.2	3.2
LI	36	70.3	1.4	64.1	1.8	74.4	1.2	70.0	2.1
BI	36	72.5	1.6	67.3	1.5	73.1	1.2	69.2	2.1
LG	36	71.7	1.1	66.9	2.0	74.6	1.1	70.2	1.9

4.3. Fundamental Frequency

F0 values at the onset of the vowel *ao* were extracted and for three out of four subjects, F0 is significantly higher when the syllable *lao* follows unstressed syllables than when it follows stressed ones. For example, for subject LI, with a mean F0 of 190 Hz and a standard deviation of 16 Hz, the starting F0 is significantly higher when vowels follow unstressed syllables than it is when they follow stressed ones, with a mean F0 of 164 Hz and a standard deviation of 21 Hz ($t(70)=5.992$, $p<0.00001$).

4.4. Discussion II

The results once again show an interaction between consecutive syllables as shown in the difference in stress status. The carryover effect, caused by preceding syllables, leads to a significant difference in segment intensity, and probably some in F0, but very little difference in segment duration. Furthermore, the effect on vowel intensity gradually decreases, suggesting this interaction to be some sort of localized phenomenon which is probably associated with a natural articulatory compensation effect — segments have a lesser intensity when following stressed syllables and a greater intensity when they follow

unstressed syllables. However, this also produces a tendency to form alternating strong and weak syllables.

5. CARRYOVER AND ANTICIPATORY EFFECTS

The previous discussion has demonstrated that interaction between successive syllables involves both carryover and anticipatory effects, and follows a general rule that unstressed syllables cause neighboring syllables to become stronger, while stressed syllables cause neighboring syllables to become weaker, in terms of variations in duration, intensity and, quite likely, F0. One of implications that can be drawn from this interaction is that the prominence of strong syllables is not only established by their lexically stressed status, but is also influenced by neighboring syllables. In this sense, syllables in Mandarin speech are phonetically realized through the tendency to form alternating strong and weak syllables. In the following section, we will examine the differences in carryover and anticipatory effects, specifically focusing on segment duration and intensity.

5.1. Segment Duration

The average durations of the consonant *ch* and vowel *i* in the syllable *chi* were calculated in four sentences: *lao3ren2 chikui1*, *lao3po chikui1*, *lao3ren2 chille*, and *lao3po chille*. Both carryover and anticipatory effects can be seen. The following table only provides the data for subject LI. It is clear that the anticipatory effect is stronger than the carryover one. The results from a two-factor ANOVA test, which was used to test the interaction between the two effects, supports this.

	Closure	Release	Consonant	Vowel
SS/SS (6)	34	100	133	64
SU/SS (6)	54	85	138	71
SS/SU (6)	37	100	159	90
SU/SU (6)	31	106	162	109

5.2. Intensity

It can be seen, by observing the overall contours of intensity, that the effect on intensity is localized, an apparent carryover effect found in the initial part of vowels, while the anticipatory effect is more pronounced in the final part of vowels. Moreover, the carryover effect is slightly stronger (or more manifest) than the anticipatory effect in general, which is in contrast to the effects on duration. The following table shows variations in intensity at the middle of consonant *l*, the onset of vowel *ao*, and the middle of vowel for subject WJ and LI.

	WJ's data			LI's data		
	middle C	onset V	middle V	middle C	onset V	middle V
SS/SS	62.5	64.8	55.9	64.3	68.7	64.4
SU/SS	69.0	72.9	58.3	70.5	74.2	69.3
SS/SU	62.3	64.7	47.9	63.7	70.2	58.4
SU/SU	71.1	71.7	55.2	70.4	73.3	65.5

6. FINAL DISCUSSION

The acoustic results in this study reveal that there is a strong interaction between successive syllables. Unstressed syllables tend to strengthen neighboring stressed syllables in terms of longer duration, greater intensity and probably higher pitch, while stressed syllables tend to weaken them. The interaction may occur within a polysyllabic prosodic constituent, realized by the anticipatory effect, or across the boundaries of prosodic constituents, realized by the carryover effect. Our results show that the anticipatory effect strongly affects vowel duration but not so much intensity, while the carryover effect affects the intensity of subsequent consonants and vowels, but has very little influence on duration.

The view that prominence in syllables with normal stress is not equal is shared by many researchers. However, there are a number of questions that are yet to be answered. Firstly, can the acoustic differences found in this study be used to classify (phonetic) stress? For example, stressed syllables in SU are defined as main or primary stress (S1), while initially stressed syllables in SS are termed secondary or weak stress (S2). However, if we were to treat the differences in acoustic variables as separate (phonetic) stresses, as opposed to an interaction, we would be left with a difficult dilemma. For example, how do we assign different degrees of stress to the syllable *lao*, when it follows either stressed or unstressed syllables? Hence, it seems better to treat them syntagmatically rather than paradigmatically. In this sense, the notion of two degrees of stress (at the lexical level), as proposed by Chao [1], is suitable for Mandarin, and other differences in prominence can be treated as the context dependent variations. Secondly, what kind of factors affect the relative prominence of normal stressed syllables? Does spoken Mandarin have fixed or predictable stress patterns? Chao has proposed a position-determined notion, which is commonly accepted in Chinese linguistics. However, there is another view, proposed by Kratochvil, which has not been given sufficient scrutiny. According to Kratochvil, there are no stable (predictable) stress patterns for Mandarin polysyllabic words, and any perceived stress patterns emerge from the tendency to alternate strong and weak syllables [4, 5]. The results from this study support Kratochvil's view. Thirdly, can the results from this study be used to create a metrical framework for Mandarin? If we take the interaction between successive syllables as a general tendency that leads to some stressed syllables becoming weaker and others stronger, then we could tentatively construct grids for Mandarin polysyllabic prosodic constituents within a framework of metrical phonology [9]. In the following table, the first level is the syllable level where each syllable is assigned with an **x**. The second shows the lexically stressed and unstressed syllables, with only normal stressed syllables being assigned **x**, while the third shows the potential relative prominence caused by that interaction, with syllables of greater prominence being assigned **x**.

(1) SS/SS	(2) SS/SU	(3) SU/SS	(4) SU/SU
x x	x x	x x	x x
x x x x	x x x	x x x	x x
x x x x	x x x x	x x x x	x x x x

The stress patterns are fairly clear in each of the above four contexts, except for (3), where either the ultimate or the penultimate syllable is

stronger in terms of the combined effects from the preceding U and ensuing S syllables. It would be fair to say, however, that the final syllable tends to be stronger when the anticipatory effects are stronger than the carryover ones, particularly when the later effect occurs within a disyllabic constituent. Therefore, two patterns, **SWWS** and **SWSW**, most likely occur in a 2/2 constituent. The same rule could be applied in Mandarin trisyllabic, quadrisyllabic, possibly even five-syllable prosodic constituents. In phonological terms, the patterns demonstrate a tendency to produce alternating strong and weak syllables, in the sense of allowing a kind of stress lapse, but not a stress clash [10].

This study provides acoustic evidence which supports the notion that Mandarin speech has unstable stress patterns and that the relative prominence of consecutive syllables is determined partly by their lexical stress status and partly by the interaction whereby unstressed syllables make neighboring syllables more prominent and stressed syllables make neighboring syllables less prominent, thereby engendering a pattern of alternating strong and weak syllables. This study not only has implications for our understanding of Mandarin stress but also provides insights into Mandarin rhythm. Since the task of approaching stress and rhythm in a language must always involve both speech production and perception, a further study involving perceptual experiments to determine whether differences in acoustic variables contribute to the perception of prominence must also be carried out.

7. REFERENCES

1. Chao, Y. R. *A Grammar of Spoken Chinese*, University of California Press, 1968.
2. Fear, B., Cutler, A. and Butterfield, S. "The strong/weak syllable distinction in English," *Journal of the Acoustical Society of America* 97, 1893-1904, 1995.
3. Harrington, J., Cassidy, S., Fletcher, J. and McVeigh, A. "The mu+ system for corpus based speech research," *Computer Speech and Language* 7, 305-331, 1993.
4. Kratochvil, P. "On the phonology of Peking stress," *Transactions of the Philological Society*, 154-178, 1968.
5. Kratochvil, P. "Stress shift mechanism and its role in Peking dialect," *Modern Asian Studies*, 8:433-458, 1974.
6. Kratochvil, P. "Beijingshua zhengchang huayu li de qingsheng," *Zhongguo Yuwen* 5, 330-345, 1987.
7. Li, W. M. "Shilun qingsheng he zhongyin," *Zhongguo Yuwen* 1, 35-40, 1981.
8. Lin, M. C. Yan, J. Z. and Sun, G. H. "Beijingshua liangzizu zhengchang zhongyin de chubu shiyan," *Fanyan* 1, 57-73, 1984.
9. Nespors, M. and Vogel, I. *Prosodic Phonology*, Dordrecht: Foris, 1986.
10. Nespors, M. and Vogel, I. "On clashes and lapses," *Phonology* 6, 69-116, 1989.
11. Tseng, C. Y. "Some stress related acoustic features of disyllabic words in Mandarin Chinese," *Bulletin of Institute of History and Phonology* LIX, 313-351, 1988.

