

A Study on the Temporal Realization of Segments under the Effect of Speaking Rate and Focus

Maolin Wang, Guangling Zi

^{1,2} (College of Chinese Language and Culture, Jinan University, Guangzhou, China)

ABSTRACT : *In this study the temporal effect of speaking rate and focus is analyzed. The test syllables are /pi, pa, pu/, which are embedded in disyllabic words, and the subjects are eight native speakers of standard Chinese. Consonantal VOT and vowel duration, together with duration of the preceding and following segments, are analyzed. It is found that the effect of speaking rate is significant for vowel duration and consonantal VOT of the key syllable, and for the duration of the preceding and following segments. The effect of focus is also significant. Focal lengthening is obvious for vowel duration and consonantal VOT, and it can also spill out to the following segments.*

KEYWORDS : *Duration, focus, speaking rate, VOT, vowel*

I. INTRODUCTION

This study investigates the effect of speaking rate and focus on the segment duration. The temporal realization of segments and syllables is basically determined by their inherent durations, however, actual segmental duration is also affected by a number of other factors, such as speaking rate, focus, and segmental environment. It has been shown that speaking rate is one of the crucial determinants of actual segmental duration, and that as speaking rate slows, VOT and syllable duration increase systematically [1]. Miller et al. [2] investigated the VOT of voiced and voiceless stops at eight rates of speech. Considering the mean VOT values as a function of syllable duration, they found that as overall syllable duration increased there was a considerable increase in VOT values for voiceless stop, but only a minimal increase in VOT values for voiced stop. That is, for both voiced and voiceless stop consonants, syllable duration increased as speaking rate slowed, but VOT values for voiced and voiceless stops were affected asymmetrically. Thus, speaking rate had a greater effect on VOT values for voiceless stops than for voiced stops.

Crystal and House [3] claimed that in their speakers' reading of an English passage, the variation of articulation rate, measured across interpause stretches, was a function of the phonological characteristics of the interpause stretch, specifically, the average number of phones per syllable and the proportion of either stressed phones or stressed syllables in the interpause stretch. In a similar study of Dutch, Eefting [4] used stretches produced in their original context as opposed to stretches embedded into a carrier phrase, and found a significant correlation between articulation rate and the phonological characteristics of the stretch only when the original context was absent. Eefting suggested that variation in articulation rate is not determined solely by phonological factors but reflects also the intention of the speaker to accommodate the audience, i. e., to facilitate the communication process.

Focus refers to syllables or words that are emphasized by the speaker to indicate that they are new or otherwise informative to the listener. In languages such as English, Dutch and Swedish, the primary correlate of focus is a pitch accent associated with the prosodic head of the focus domain. This pitch accent is usually also accompanied by other correlates such as longer durations within the word carrying the accent [5]. In English, e.g., when a word is focused, all syllables within the word are lengthened. Furthermore, even when only one syllable out of the word is focused,

the whole word lengthens [6].de Jong [7] studied the effects of stress and focus on vowel duration and quality in English. She finds that there are durational differences between vowel categories and between vowels preceding voiced and voiceless stops, and there is an effect of stress and focus on the size of voicing-induced lengthening of vowels preceding stop consonants. Focus effects are mediated by stress such that increases in durational differences are localized largely in syllables which are primary stressed and accented.

The effects of focus and stress are found to be largely the same, and they can be said to be parallel physical effects. For Chinese, Chen [8] examined the patterns of durational adjustment of mono-morphemic four-syllable words when different constituents of the word are focused for correction. She found that when the word is in utterance medial position, focus induces evident lengthening. When a focused domain is multi-syllabic, the distribution of lengthening is non-uniform: there is a strong tendency of edge effect with the last syllable lengthened the most, and there is also spill-over lengthening on the neighboring syllables outside the focused constituent.

Much research work has been done on the realization of pitch and duration of focus in Chinese. It is shown that focus patterns are implemented as pitch range variations imposed on different regions of an utterance. The pitch range of tonal contours directly under focus is substantially expanded; the pitch range after the focus is severely suppressed; and the pitch range before the focus does not deviate much from the neutral-focus condition. Thus, there seem to be three distinct focus-related pitch ranges: expanded in non-final focused words, suppressed in post-focus words, and neutral in all other words. It is also shown that the on-focus force increases the rising slope of the rising tone in Chinese, and research on focus in both English and Chinese has shown many similarities between the two languages [9, 10]. This paper aims at probing the temporal effect of two factors in Chinese, i.e., speaking rate and focus. To be specific, the effect of the factors on VOT, vowel duration, as well as durations of segments preceding and following the key syllable, will be investigated.

II. METHOD

2.1. Subjects, stimuli and recording

Eight native speakers of Standard Chinese, four males and four females, participated in this experiment, and the test syllables are /pi, pa, pu/, which are embedded in disyllabic words, i.e., ‘pixie’ (leather shoes), ‘pahao’ (braised oyster) and ‘puxi’ (cattail mat). These words are included in sentences with which the two critical factors are manipulated: speaking rate and focus. The two speaking rates are normal and slow rates, and the two focus conditions are focused and unfocused. Sentences for /pa/ are listed as follow,

(a) focused

Q: Shaohao he lamian yiyang gui ma?

Is roast oyster as expensive as ramen?

A: Bu, genju diaocha, **pahao** he lamian yiyang gui.

No, according to the survey, braised oyster is as expensive as ramen.

(b) unfocused

Q: Pahao he lamian yiyang gui ma?

Is braised oyster as expensive as ramen?

A: Bu, genju diaocha, pahao he **lumian** yiyang gui.

No, according to the survey, braised oyster is as expensive as braised noodles.

Recording was done in a sound-proofed recording room, and the sentences were randomized for each repetition of each subject. The questions were recorded in advance by an experienced speaker in two rates, normal and slow rate, and they were played from the loudspeakers to the subject. Before recording, the subject was allowed to browse the sentences in order to be familiar with them. The subject was asked to read the answers in the same rate as the questions. To improve statistical reliability, the entire corpus was repeated four times, with 2 minutes’ rest after each time, and in total 384 sentences (2 rates \times 2 focus conditions \times 3 vowels \times 4 repetitions \times 8 speakers) were collected and analyzed in the present study.

2.2. Measurements

Duration of the preboundary vowel: The duration of the preboundary vowel, which is always /a/ as in ‘diaocha’ (survey).

Voice onset time (VOT): VOT of the aspirated stop /p/ of the key syllable /pi, pa, pu/.

Duration of the vowel of the key syllable: The duration of the vowel in the key syllable /pi, pa, pu/.

Duration of the onset and the rhyme of the following syllable: In Chinese, most of the syllables are composed of two parts, the onset and the rhyme, except the ‘zero-onset’ syllables. For example, in the syllable of ‘hao’, the onset is ‘h’ and the rhyme is ‘ao’. But in zero-onset syllable like ‘ai’, there is no onset, only the rhyme ‘ai’.

In the present study, the duration of the onset and the rhyme of the syllable following the key one, like ‘hao’ in ‘pahao’, will be investigated. Annotation was done by an experienced annotator, and repeated measures ANOVA analyses were done for the statistical evaluation of the influence of the factors of speaking rate and focus.

III. RESULTS

3.1. Duration of the preboundary vowel

Fig. 1 shows the duration of the preboundary vowel, broken down by speaking rate and focus, from which it can be seen that the effect of speaking rate is obvious. It is shown from repeated measures ANOVA results that the effect of both factors is significant, speaking rate: $F(1, 95) = 288.2, p < 0.001$; focus: $F(1, 95) = 4.79, p = 0.031$. The duration of the preboundary vowel is comparatively long in slow speaking rate, at pre-nonfocused position.

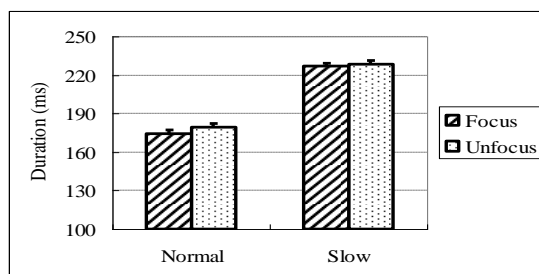


Fig. 1 Duration of the preboundary vowel

Further observation shows that there is no significant speaking rate \times focus interaction: $F(1, 95) = 0.775, p = 0.381$. The effect of speaking rate and focus on the duration of the preboundary vowel is consistent on different levels of the two factors.

3.2. VOT of /p/ in the key syllable

Fig. 2 graphs the VOT of the aspirated stop /p/ of the key syllable. Repeated measures ANOVA result shows that the effect of speaking rate and focus is significant, speaking rate: $F(1, 95) = 67.9, p < 0.001$; focus: $F(1, 95) = 83.5, p < 0.001$, with VOT of slow rate longer than normal rate, and that of focused condition longer than that of unfocused one.

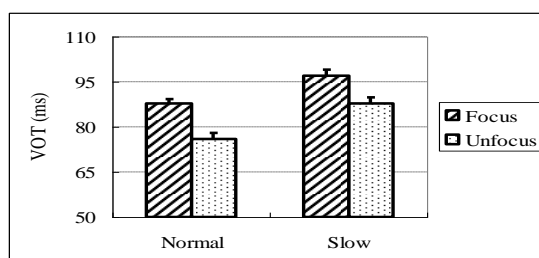


Fig. 2 VOT of /p/ in the key syllable

Regarding interactive effect, similar to that on the duration of the preboundary vowel, it is shown that the interaction of speaking rate \times focus is not significant: $F(1, 95) = 1.88, p = 0.173$. The effect of speaking rate and focus on the VOT of the consonant /p/ is consistent on different levels of the two factors.

3.3. Duration of the vowel of the key syllable

Fig. 3 shows the duration of the vowel of the key syllable, and it is shown from repeated measures ANOVA results that the effects of both factors are significant, speaking rate: $F(1, 95) = 893.5, p < 0.001$; focus: $F(1, 95) = 369.1, p < 0.001$. The duration of the vowel is comparatively long in slow speaking rate and under focused condition.

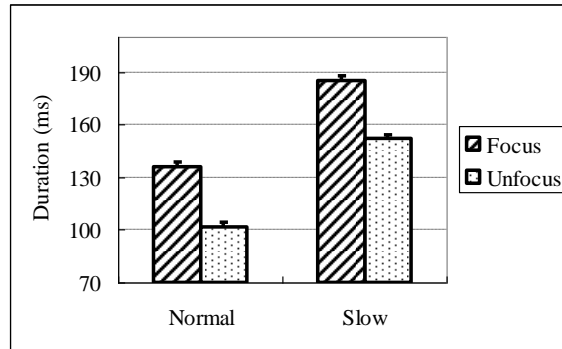


Fig 3. Duration of the vowel of the key syllable

As for interactive effect, the interaction of speaking rate \times focus is not significant: $F(1, 95) = 0.409, p = 0.524$. There is no interactive effect for the two factors on the duration vowel of the key syllable.

3.4. Onset duration of the following syllable

Fig. 4 graphs the duration of onset of the following syllable. Repeated measures ANOVA result shows that the effect of both of the two factors is significant, speaking rate: $F(1, 95) = 231.5, p < 0.001$; focus: $F(1, 95) = 60.1, p < 0.001$. The duration of the onset is comparatively long in slow speaking rate, and following focused syllable.

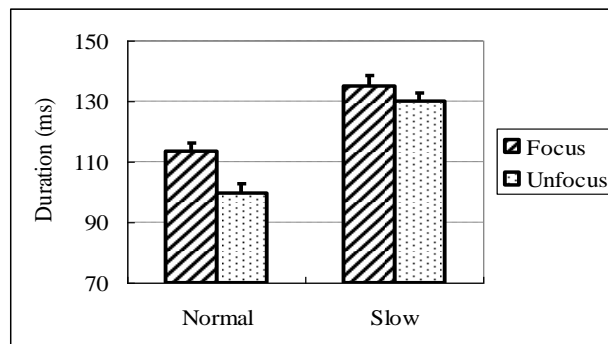


Fig. 4. Onset duration of the following syllable

In regard to interactive effect, it is shown that the interaction of speaking rate \times focus is significant: $F(1, 95) = 16.1, p < 0.001$, which is due to the disproportionate effect of focus in the two speaking rates. The effects of the two speaking rates will be elaborated below.

(1) In normal speaking rate, the effect of focus is significant: $F(1, 95) = 80.1, p < 0.001$, with the duration of the onset under focused condition longer than that under unfocused condition.

(2) In slow rate, the effect of focus is also significant, only in that the significant level is low: $F(1, 95) = 9.25, p = 0.003$.

3.5. Rhyme duration of the following syllable

Fig. 5 displays the duration of the rhyme of the following syllable, and it is shown from repeated measures ANOVA result that the effects of both of the two factors are significant, speaking rate: $F(1, 95) = 528.1, p < 0.001$; focus: $F(1, 95) = 40.8, p < 0.001$. The duration of the rhyme is comparatively long in slow speaking rate, and following focused syllable.

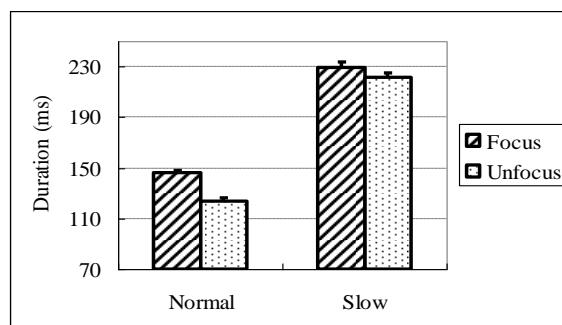


Fig. 5. Rhyme duration of the following syllable

Regarding interactive effect, the interaction of speaking rate \times focus is significant: $F(1, 95) = 14.4, p < 0.001$, which is also due to the disproportionate effect of focus in the two speaking rates. The effects of the two speaking rates will be elaborated below.

(1) In normal speaking rate, the effect of focus is significant: $F(1, 95) = 54.1, p < 0.001$, with the duration of the onset under focused condition longer than that under unfocused condition.

(2) In slow rate, the effect of focus is also significant, but the significant level is low: $F(1, 95) = 7.57, p = 0.007$.

IV. DISCUSSION

4.1. Vowel duration of the key syllable

Results from the previous section show that the duration of the vowel of the key syllable is affected by both of the two factors, speaking rate and focus. Its duration is relatively long in slow speaking rate and under focused condition. This means that in Chinese, vowel is more likely to be affected by the various factors. In Chinese, there is no contrast of long vowels and short vowels, therefore, the temporal realization of vowels may be more flexible, and their durations are more sensitive to factors like speaking rate and focus. However, it can be seen that the effect of the two factors is not equal: that of speaking rate is great, and that of focus is small.

4.2. VOT of /p/ in the key syllable

VOT of /p/ is also affected by the two factors. From the result of the previous section, it can be seen that the effect of speaking rate and focus on VOT is significant, with that of slow rate longer than normal rate, and focused condition longer than unfocused one. However, compared to that on the duration of the vowel, the effect on the VOT of /p/ is small.

4.3. The preceding and the following segments

Regarding the duration of the preboundary vowel, the effects of speaking rate and focus are significant. The duration is comparatively long in slow speaking rate and at pre-nonfocused position. This is a little strange: For the vowel of the key syllable, its duration prolongs when it is under focus condition. However, for the preboundary vowel, its duration prolongs under the non-focused condition. We speculate that this is because of the isochrony principle [11]. It is generally approved that languages can be classified typologically into 'stress-timed' languages and 'syllable-timed' ones [12]. Languages like English and German are of the first type, and Chinese belongs to the second type. It is also believed that languages observe the isochrony principle. For stress-timed languages, intervals between the stressed syllables tend to be equal, and for syllable-timed languages, syllables tend to be of equal length. Results from the previous section show that the effect of focus on both the duration of the vowel and VOT of consonant /p/ is significant. Under focused condition, the duration of the vowel, as well as the VOT of the consonant, is greatly lengthened. Chinese is a syllable-timed language, so syllables tend to be of same duration.

With the effect of the isochrony principle, under focused condition, since the duration of the key syllable prolongs, that of the preceding vowel will be reduced. Therefore, the duration of the preboundary vowel is long under non-focused condition. As for the duration of the onset and the rhyme of the following syllable, it is shown that the effects of both of the two factors are significant, with durations of onset and rhyme comparatively long in slow speaking rate, and following focused syllable. That is, focus lengthening spills out of the focused syllable to the following syllable, not only to the onset, but also to the rhyme. In Chinese, a word is

psychological unit, which is semantically as a whole. When a syllable in the word is focused and lengthened, the other syllable will also be influenced and lengthened.

V. CONCLUSION

In this study, the temporal effect of speaking rate and focus is analyzed. It is found that the effect of speaking rate is significant for vowel duration and consonantal VOT of the key syllable, and for the duration of the preceding and following segments. The effect of focus is also significant. Focal lengthening is obvious for vowel duration and consonantal VOT, and it can also spill out to the following segments. In Chinese, there is no phonological contrast of long vowels and short vowels, therefore, the temporal realization of vowels may be more flexible, and their durations are more sensitive to factors like speaking rate and focus. Chinese is a syllable-timed language, so syllables tend to be of same duration. With the effect of the isochrony principle, under focused condition, since the duration of the key syllable prolongs, that of the preceding vowel will be reduced. Therefore, the duration of the preboundary vowel is long under non-focused condition.

VI. ACKNOWLEDGEMENTS

This work was supported in part by the Social Science Research Project of Guangdong Province, Grant No. GD11CWW04, as well as the Humanity and Social Science Research Project for Colleges in Guangdong Province, Grand No. 11WYXM012.

REFERENCES

- [1] L. E. Volaitis and J. L. Miller, "Phonetic prototypes: influence of place of articulation and speaking rate on the internal structure of voicing categories," *Journal of the Acoustical Society of America*, 92, pp. 723-735, 1992.
- [2] Miller, J. L., Green, K. P. and Reeves, A. "Speaking rate and segments: a look at the relation between speech production and speech perception for the voicing contrast," *Phonetica*, 43, 106-115, 1986.
- [3] Crystal, T. H. and House, A. S. "Articulation rate and the duration of syllables and stress groups in connected speech," *Journal of the Acoustical Society of America*, 88(1), 101-12, 1990.
- [4] Eefting, W. "Production and perception of temporal variation: an explorative study," In *OTS Yearbook*, pp. 13-38. Research Institute for Language and Speech: University of Utrecht, 1990.
- [5] M. Heldner and E. Strangert, "Temporal effects of focus in Swedish," *Journal of Phonetics*, 29, pp. 329-361, 2001.
- [6] T. Cambier-Langeveld and A. Turk, "A cross-linguistic study of accentual lengthening: Dutch vs. English," *Journal of Phonetics*, 27, pp. 255-280, 1999.
- [7] K. de Jong, "Stress, lexical focus, and segmental focus in English: patterns of variation in vowel duration," *Journal of Phonetics*, 32: 493-516, 2004.
- [8] Y. Chen, "Durational adjustment under corrective focus in Standard Chinese," *Journal of Phonetics*, 34, pp. 176-201, 2006.
- [9] Y. Xu and C. Xu, "Phonetic realization of focus in English declarative intonation," *Journal of Phonetics*, 33, pp. 159-197, 2005.
- [10] Y. Xu, "Effects of tone and focus on the formation and alignment of F0 contours," *Journal of Phonetics*, 27, pp. 55-105, 1999.
- [11] J. Laver, *Principles of phonetics*, Cambridge: Cambridge University Press, 1994.
- [12] K. L. Pike, *The intonation of American English*, Ann Arbor: University of Michigan Press, 1945.