The durational patterns of syllables in Standard Chinese

Cao, Jianfen

Institute of Linguistics
Chinese Academy of Social Sciences
Beijing 100732, CHINA

ABSTRACT

The purpose of This study is trying to search the durational patterns of syllables in Standard Chinese by observing their variations in different contexts. For example, durational patterns may be closely related to syntactic and semantic constraint of sentences. Thus, the test sentences were designed. In addition, for the convenience of discussion, the durational measurement is also extended to the acoustic record of monosyllabic and bisyllabic words. Therefore, the durational realization of a syllable in larger structures may be not only determined by their position in the sequence, but also by the morphological constraint of the structure. To simplify our analysis, we just take trisyllabic word as the test example. Secondly, syllable as a basic conveyer for the syntactic and semantic information, its durational variations may be closely related to syntactic and semantic constraint of sentences. Thus, the test sentences were designed. In addition, for the convenience of discussion, the durational measurement is also extended to the acoustic record of monosyllabic and bisyllabic words which were made in our previous investigations [17, 21].

1. INTRODUCTION

Individually, segments or syllables' duration in natural speech has wide range of elasticity, it is typically true in Standard Chinese (SC). Nevertheless, the speech tempo is still kept in a regular range, it seems that there may exist some relative invariance in temporal distribution, by which the speech tempo is controlled.

Considering of the importance of the syllable in SC, the present investigation is concentrate on the control of syllable duration. Against this target, two sets of materials were tested in this study. The first set consists of 14 pairs of trisyllabic words, and the second set contains 6 pairs of sentences, all of them were designed according to the background as follows. Firstly, syllable, as one of the basic units in Chinese speech, is usually running in some relatively fixed groups, among them, the minimal one is bisyllabic word. This kind of words inherently includes two types of stress and the corresponding temporal patterns [21]. Moreover, the bisyllabic words in SC constructionally has relative stability and strong power in constituting new words, so all of the other polysyllabic structures in SC are essentially different combinations built from monosyllabic and bisyllabic words. Therefore, the durational realization of a syllable in larger structures may be not only determined by their position in the sequence, but also by the morphological constraint of the structure. To simplify our analysis, we just take trisyllabic word as the test example. Secondly, syllable as a basic conveyer for the syntactic and semantic information, its durational variations may be closely related to syntactic and semantic constraint of sentences. Thus, the test sentences were designed. In addition, for the convenience of discussion, the durational measurement is also extended to the acoustic record of monosyllabic and bisyllabic words which were made in our previous investigations [17, 21].

Figure 1. Durational patterns of syllables summarized from various isolated polysyllabic words:

- A. NM type bisyll. word; A(1) 1-2 NM type trisyll. word;
  A(2) 1-2 NM type trisyll. word.
- B. NT type bisyll. word; B(1) 1-2 NT type trisyll. word;
  B(2) 1-3 NT type trisyll. word.

Table 1. Average duration (ms) of different variations for syllables in moderate speech and 6-7 for faster speech. This tendency seems not to be an accidental event, some previous studies have provided evidence from either production or perception.

<table>
<thead>
<tr>
<th>Position</th>
<th>In moderate speech</th>
<th>In faster speech</th>
</tr>
</thead>
<tbody>
<tr>
<td>Final</td>
<td>332.5 357.5 390</td>
<td>245 290 260</td>
</tr>
<tr>
<td>Initial</td>
<td>260.7 268.8 390</td>
<td>222.5 263.3</td>
</tr>
<tr>
<td>Vocalic</td>
<td><em>(a)</em> 193 175</td>
<td>185 135</td>
</tr>
<tr>
<td><em>B</em></td>
<td>187 148</td>
<td>161.3 122.5</td>
</tr>
</tbody>
</table>

*(a)* 2nd syllables in trisyllabic words;
*(b)* 3rd syllables in quadsyllabic words.

Table 1. Average duration (ms) of different variations for syllables in moderate speech and 6-7 for faster speech. This tendency seems not to be an accidental event, some previous studies have provided evidence from either production or perception. Based on these findings coming from different languages, we would suppose that the marginal seven or less of syllable number uttered per second might be one of the relative invariance in speech tempo existed throughout the languages in the world. Thus, it may be served as a relative scale for the comparison in this aspect.

2.2. Inherent difference

Some inherent difference of syllable duration in SC is observed from the isolated syllables. The main control factors for this difference, as it is universally acted in many languages, includes articulatory manner of the initial consonants and the opening of component vowels. In addition to these in SC, this kind of difference is also closely related to the tonal distinction and stress type. Generally, a syllable with the third tone is usually longer than that with other tones ceteris paribus; and a syllable with normal (NH) type stress is clearly longer than that with neutral (NT) type's. Some previous studies have reported that the duration ratio of NT to NM type syllables is about 1:2 [51] to 3:5 [11], and it is contextually dependent, when neutral type bisyllabic words are uttered in connected speech, the durational difference between NT and NM type syllables becomes even more sharpened [21].

2.3. Contextual variability

2.3.1. Variations in isolated words

Figure 1. gives specific variations of syllable in isolated polysyllabic words. The dark bar on the far left in the figure represent a relative scale for the comparison among the variations, its value is taken from...
the average duration of the word-initial syllables; The bars next to the dark one represent those patterns summarized from bisyllabic words tested in our previous investigation (2), and all of the rest are the patterns occurred in different types of trisyllabic words. Here the bars with oblique lines are the NN type tokens, and those with cross lines are the NT type tokens. In each case, the shadowed part of the bar shows a scope that a syllable's duration is exactly likely reached, and the part within dotted lines illustrates the extent that possibly reached.

From Figure 1, some effects could be clearly observed. (a)Position effect: Syllable duration in SC does follow position-effect rule, which is universally acted in many languages, such as word-initial shortening, word-final lengthening, and the intervocalic ones are often shorter in average. This effect also can be seen from Table 1; (b)Syntactic constraint: As one of the basic units in SC, the durational patterns of bisyllabic words are often appeared in global. Look at the cases of B(1), B(2)and B(3) in Figure 1, it's obvious that the long-short duration model for NT type's bisyllabic structure is kept in any position within the trisyllabic words, so the second syllable in B(1) and the first syllable in B(2) and B(3) have not been shortened as expected, but obviously lengthened instead, though all of them occur at the intervocalic position or word-initial position. This phenomenon illustrates that the power of position-effect is limited, and it must be played under the control of morphological restriction. 2.3.2 Variations in connected speech (a) Syntactic effect The influence comes from syntactic aspect has been observed in our previous study (2). It may be helpful to present an example here: when the test word in a sentence is followed by an independent syntactic element, the average duration of last syllable in these words is around 292 ms; while it is compressed to 254 ms when it is followed by an auxiliary word, since this kind of words is usually not an independent element and has to be attached to its previous words, consequently, results in a limitation to the last syllable of the test words.

(b) Units' length and speech rate If we look at back to the situation shown in Table 1, and make a simple comparison between the first row and the last two rows respectively or that between last two rows, it can be seen that syllable duration does decrease with the increase of syllable number in certain unit. At the same time, the relationship on inverse proportion between syllable duration and speech rate is also found from the cases in moderate' column and 'faster' column in Table. However, that is not a simple and linear relation, because speech tempo is also adjusted by presence / absence and longer / shorter of the possible pauses occurred at phrase boundary, and syllable's elongation or compression performed in either case is only a tendency in gross. (c) Semantic influence Speech sound is the surface output represented the underlying semantic input, so the surface realization of a syllable duration in real speech is reasonably related to semantic content. Our observation indicates that there is usually carried out through the difference in accent stress. Figure 2 gives one of the examples, in which, four different statements are made from the same syllable sequence but represent different semantic contexts, and results in different distribution of syllable duration.

3. CONCLUSION
Based on the limited observation taken in this study, the durational patterns of syllables in Standard Chinese could be summarized as follows:
(a) Syllables as monosyllabic words uttered moderately in isolation are around 500 ms. Some intrinsic difference is existed among them, and it is conditioned by some inherent factors related to syllable itself. However, all of them will be shortened when they occur in connected speech, the average ratio between these two cases is around 100:70.
(b) The durational realizations for individual syllable in real Chinese speech have a wide dynamic range, the lower threshold is around 100 ms, and the higher one's is about 600 ms. The contextual variations are shown in Table 1 of the text.
(c) The durational patterns summarized above in SC is governed by multiple factors acted in the different levels. The main factors, besides those inherent ones related to syllable itself, are coming from speech rate, syllable's position in the utterance, units' length in which the syllable appeared, morphological constraint of speech unit, difference in syntactic contexts, and the semantic requirement of certain sentences. Moreover, it seems that the last factor is the most active and powerful one, all of the others must be acted under its control. Therefore, these factors are hierarchically worked on a top-down way, the elongation or compression for a syllable is performed non-linearly through the mechanisms of compensation and adjustment, rather than in average or complete proportion.

4. REFERENCES