Working Papers 25 1983
Linguistics-Phonetics
Lund University, 105-111

DURATION OF STANDARD CHINESE WORD TONES IN DIFFERENT SENTENCE ENVIRONMENTS

Magnus Nordenhake and Jan-Olof Svantesson

1. BACKGROUND

In Standard Chinese (pǔtōnghuà) there are four tones: 1. high (denoted ¯); 2. rising (´); 3. low (dipping, ˇ) and 4. falling (`). Apart from fundamental frequency, two other acoustic correlates have been deemed important in the manifestation of the Standard Chinese word tones: intensity and duration.

Kratochvíl (1968) refers to tones produced in isolation when he describes the length of the first tone as "slightly above average"; the second tone as "slightly below average"; the third tone as "well above average"; and the fourth tone as "far below average". As far as we are aware, there is no published investigation of the duration of Standard Chinese tones in controlled sentence environments.

In this paper the duration of Standard Chinese tones in different environments (sentence medial and final), and under different focus conditions (focus or non-focus) is investigated. It is shown that the durations of the four tones are affected differently by different environments. In particular, the fourth tone shows the shortest duration of the four tones in sentence final position, but is the longest tone in sentence medial position. All four tones are lengthened in focus position.

2. PROCEDURE

The speaker, a male resident of Běijīng in his forties, was asked to read at normal speed a number of questions and answers at the sound treated studio at the Institute of Phonetics in Lund. The material was recorded on a Studer tape recorder.

The analyzed sentences were of the type (the standard pīnyīn transcription is used):

Wǒmen V le Wáng N 'We V-d Wáng N'

Here, V is a monosyllabic verb, and N is a given name. The studied syllables (i.e. V and N) were chosen so that all four tones were represented in both positions, while these syllables were all of the type Ca. By necessity, different consonants had to be chosen. In the analyzed material, the medial syllable V is followed by an unstressed syllable in order to minimize the influence of tone sandhi.

The total material consisted of the following sixteen sentences:

Wǒmen { shā / chá / dǎ / mà } le Wáng { Bā / Pá / Mǎ / Chà }

We { kill / examine / beat / scold } -d Wáng { Bā / Pá / Mǎ / Chà }

Each sentence was recorded twice, once with focus on the sentence medial verb, and once with focus on the final noun. This was achieved by recording each sentence as the answer to two different questions: Nǐmen gēn Wáng N zuòle shénme? 'What did you do with Wáng N?' the answer of which focussed on the sentence medial verb; and Nǐmen V le shéi? 'Who did you V?' which brought the focus to the name. The speaker read each of the 32 questions and answers twice, and wideband spectrograms were made of the answers. On the spectrograms, the durations of the vowels (a) in the studied syllables were measured.

3. RESULTS

Table 1 shows the duration of the vowel in each of the 128 analyzed syllables. On Figures 1 and 2, each point represents the average of 8 measurements with the same tone, position and focus condition. The data are arranged in different ways on the two figures in order to illustrate the effects of the factors tone, position and focus.

The results show that tones 1, 2 and 3 are longer in final position than in medial position. In fact, in our material, all

Table 1. Duration of the vowel in Standard Chinese syllables of the type Ca (ms).

	Medial		Final	
	Focus	Non-focus	Focus	Non-focus
Tone 1	166	158	205	217
	166	181	213	205
	174	150	185	205
	181	166	229	197
	158	148	221	205
	174	158	221	197
	166	150	213	197
	181	158	229	213
mean:	171	159	214	204
Tone 2	142	103	252	221
	174	142	252	205
	150	95	221	213
	174	142	237	221
	158	142	229	205
	189	150	252	221
	142	118	237	205
	166	142	237	197
mean:	162	129	240	211
Tone 3	189	158	315	237
	174	166	284	205
	189	166	292	260
	189	174	292	197
	205	174	308	252
	174	181	300	252
	197	150	308	270
	189	174	276	260
mean:	188	168	297	242
Tone 4	181	184	166	142
	197	181	158	142
	197	166	166	158
	213	189	174	181
	181	174	150	150
	197	181	174	174
	181	174	158	158
	221	189	174	110
mean:	196	180	165	152

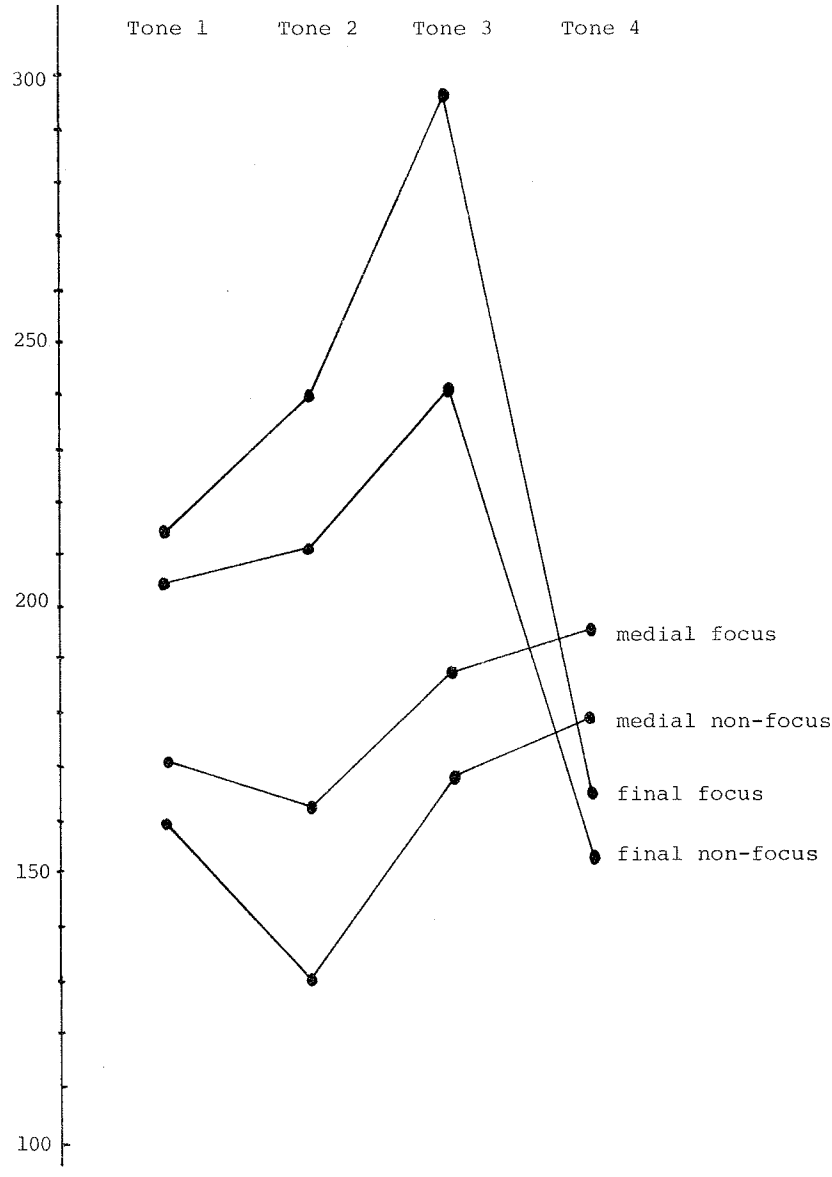

Figure 1. Average duration for each combination of position and focus conditions over the four tones.

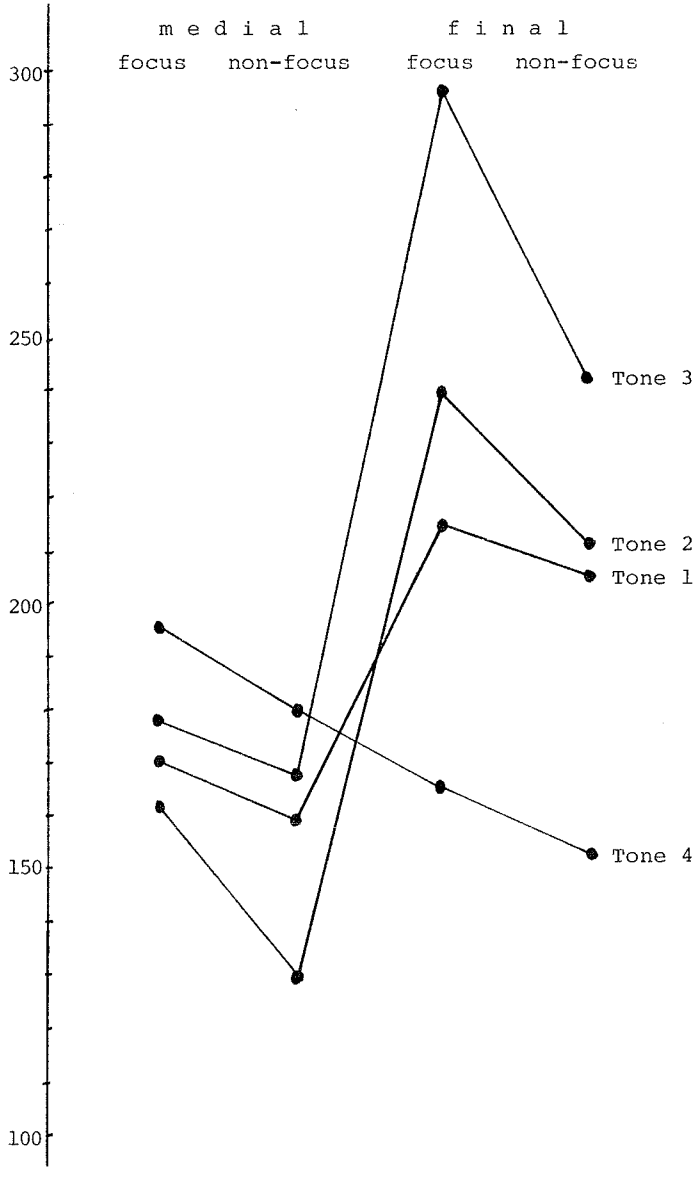

Figure 2. Average duration of the four tones over each combination of position and focus conditions.

finals are longer than all medials with the same focus condition for these three tones. Tone 4, however, is on the average shorter in final than in medial position (see Fig. 1).

Thus the factor "position" influences the different tones in different ways. The factor "focus", however, always has the same effect: a focussed vowel becomes longer irrespective of its tone.

In order to test statistically the interaction between the factors tone, position and focus, a three-way analysis of variance was performed with the following result:

Factor	Sum of squares	Degrees of freedom	Test quantity	
tone	45 565	3		
position	69 425	1		
focus	17 743	1		
tone-position	72 327	3	115.0	($p \ll 0.001$)
tone-focus	3 904	3	6.2	($p < 0.001$)
position-focus	328	1	1.6	($p > 0.05$)
tone-position-focus	2 165	3	3.4	($0.01 < p < 0.05$)
residual	23 490	112		

Thus there is no significant interaction between position and focus. The interaction between tone and focus and between tone and position are both significant, but the tone-focus interaction is much smaller than that between tone and position. The average effects of focus and of final position are (cf. also Figure 2):

	Focus	Final position (ms)
Tone 1	11.1	44.9
Tone 2	30.6	79.8
Tone 3	37.8	91.2
Tone 4	14.7	-29.4

Thus, tones 2 and 3 are lengthened more than tone 1 and 4 in focus, and as already said, the effect of final position (as compared to medial position) is to lengthen tones 1, 2 and 3, but to shorten tone 4.

Because of these interactions, it is hardly meaningful to give the average durations of the four tones without stating the

environment. In medial position the tones are ordered in increasing duration as: T2 <T1 <T3 <T4, and in final position the order is T4 <T1 <T2 <T3. The order given by Kratochvíl 1968 for citation forms (T4 <T2 <T1 <T3) is the same as that for final position, except that the order between T2 and T1 is reversed (the difference between T2 and T1 as found here is small, however).

4 DISCUSSION

One might speculate why the falling fourth tone is shortened in final position. One possible reason is that because of sentence intonation downdrift in statements (cf Gårding, Zhāng and Svantesson 1983), the starting point for a sentence-final falling tone is rather low, so that it takes comparably short time to reach the bottom of the voice register. Perception experiments by Kratochvíl 1970 show that the fourth tone is associated with shortness. He found that shortened tones, may they be tone 1, 2, 3 or 4, are almost exclusively identified as tone 4.

Tone 3 is also somewhat exceptional since it is lengthened more than the other tones both by focus and by final position. It seems that staying for some time at a low level is an essential property of this tone, and this property is strengthened by focus and by final position. The dipping citation form of this tone is obtained by combining the intrinsic low tone with sentence intonation (cf. Gårding, Zhāng and Svantesson 1983).

REFERENCES

Gårding, E., Zhāng, J and Svantesson J-O. 1983. A generative model for tone and intonation in Standard Chinese. In this volume.

Kratochvíl, P. 1968. The Chinese language today. London: Hutchinson.

Kratochvíl, P. 1970. An experiment in the perception of Peking dialect tone. In I-L Hansson, ed., A symposium on Chinese grammar. London: Curzon Press, pp. 7-31.