

# ON NEUTRAL-TONE SYLLABLES IN MANDARIN CHINESE

Cao, Jianfen

Institute of Linguistics  
Chinese Academy of Social Sciences  
5 Jianguomennei Rd.  
Beijing 100732, CHINA

## 1. INTRODUCTION

One of the characteristic and hard to be grasped features in Mandarin Chinese is the so-called neutral-tone ( hereafter NT) syllables, which is neutralized from corresponding normal stressed syllables. Perceptually, The NT syllable sounds quite weak, and must be attached to other syllable as a dependent morpheme in polysyllabic words. However, it does not equal to the unstressed syllables in common, in stead, it forms an important morphophonemic contrast to the normal type one in bisyllabic words. Consequently, it might be of benefit to Chinese speech processing to clarify the phonetic nature of the NT syllables.

Many relevant approaches from different viewpoint [1][2][3] have been contributed before, this study is concerned in their phonetic characteristics. The discussion is based on a relevant acoustic measurements and analysis [4] referred to a corpus of materials including over 200 of NT syllables, all of them are occurred as second morpheme in different category of neutral type bisyllabic words.

The hypothesis is that the perceptual feature of the NT syllables may be caused by a combination of multiple acoustic effects, in stead of any single factor. Therefore, this kind of syllables must own their special acoustic patterns.

## II. EXPERIMENTAL RESULTS AND DISCUSSION

### 2.1 The Intensity of NT syllables

Perceptually, the stress of NT syllables is weakened, so people are often used to regard that it is a result of low intensity. However, The data obtained here illustrate that the intensity of NT syllable is not necessary lower than that of the syllable with normal stress. Table I offers a comparison on average intensity between normal syllables and the NT ones obtained from a set of duplicated bisyllabic words, in which two syllables with the same phonetic structure but different stress type. The corresponding figures show clearly that the intensity of NT syllable is lower than that of the normal one in general, but the situation is just reversed when it is preceded by a third tone syllable. This tendency also can be seen from Fig 1, so it seems to be a relational invariance in this area. Therefore, the stress of NT syllable is weakened, but not simply related to its intensity.

Table I. Comparison on average intensity(dB) between normal and NT syllables in duplicated neutral type bisyllabic words.

Original tone category	Intensity of normal syllable		Intensity of NT syllable	
	value	sd	value	sd
first tone	15.20	1.95	7.46	0.96
second tone	12.77	2.99	9.87	1.35
third tone	9.38	1.38	10.05	0.94
fourth tone	15.85	0.64	7.08	1.88

### 2.2 The pitch of NT syllables

One of the most characteristic features of NT syllable is loss of its original tone pattern. Usually, people pay most of the attention to its short duration and relative pitch height, while the dynamic range of pitch contour is often ignored [5].

According to the perceptual impression, the pitch height is usually marked with the points of three or four levels [6][7], namely, the highest one occurred after a third tone syllable, then the ones occurred after the second tone and first tone syllable respectively, the lowest one is that occurred after the fourth tone syllable.

Based on the results obtained from relevant investigation, as it has been summarized in Fig.1 and Table II, the following findings have been made [4], though it is still remained as a controversial issue. First, it is obvious that there exist two different type of pitch contour among the NT syllables, and it is significantly distinguishable ( $p < 0.01$ ) determined by the tonal difference of preceding syllables, namely, a mid-level or slightly rising contour after a third tone syllable, and a mid-falling one after all of the other tone's syllables; Second, as what can be seen from Table II, the average pitch value at the onset of NT syllables in different tonal context is closed to one another, while the value at syllable offset does show a regular order, i.e, the highest value is held by the tokens occurred after the third tone syllable, and then is that of the tokens after second and first tone syllables, the lowest one is that occurred after the third tone syllables. This order is just matching to the four-level's distinction of perceptual impression mentioned above. Consequently, the dominant effects for the perceived impression of pitch height should be coming from the whole pitch contour, rather than a single point of pitch height, at least, it is more related to the later part of NT syllables, in stead of the pitch values at the onset as some approach [8] suggested. Actually,

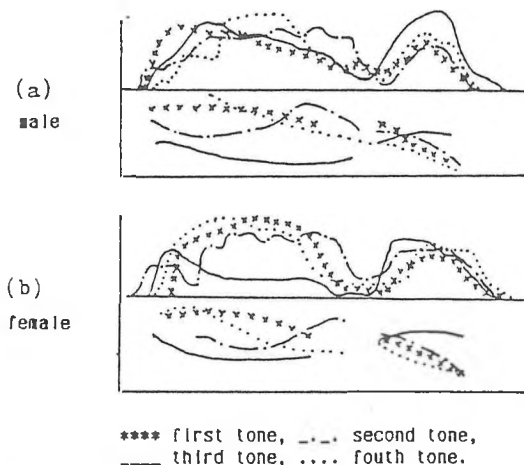


Fig.1 Comparison on intensity and pitch height between NT syllable and preceding normal syllable: The top lines in both of (a) and (b) represent amplitude contours, and the bottom ones the pitch contours.

Table II. Average value of relative pitch height(Hz) of NT syllables

preceded tona context	pitch value at the onset		pitch value at the offset	
	value	sd	value	sd
after 1st tone	152.34	21.05	89.51	11.76
after 2nd tone	174.71	11.55	108.45	27.64
after 3rd tone	136.60	6.93	146.60	3.46
after 4th tone	123.45	15.19	75.30	11.59

Table III. Comparison on duration(ms) between normal(NM) type syllables and NT syllables

original tone category	normal syllable		NT syllable		ratio of NM to NT syllables
	value	sd	value	sd	
first tone	313.21	34.33	186.76	24.78	100:60
second tone	304.92	51.14	183.02	30.50	100:60
third tone	299.80	44.89	209.13	10.56	100:70
fourth tone	302.25	60.75	192.33	28.15	100:64
plotted	305.05	5.83	192.81	9.99	100:63

this finding has been revealed by a relevant test of the synthesis[9]. It indicates that, the NT syllables do have own regular pitch patterns and it is definitely contributed to the perception of NT syllables.

### 2.3 The duration of NT syllables

Generally, the NT syllable is systematically shorter than the syllable with normal stress, the average ratio between them is about 3:5 in isolation and 1:2 in running speech[10]. Of course, it is only a rough range, the specific realization of individual tokens is more complex. First, According to the average duration (see Table III) measured from a group of duplicated bisyllabic words in isolation, the durational realization of NT syllables seems also to be determined by tonal difference of the preceding syllables. Generally, a relatively longer duration is held by the tokens occurred after third tone syllables; second, the shorter duration of NT syllable is not just caused by a simple compression of the whole syllable, but a change of temporal distribution within the syllable. Generally, in the NT syllable, a relatively greater degree of shortening is taken place in the vowel, instead of the consonant of the syllable.

### 2.4 The sound quality of NT syllables

The sound qualities in NT syllables, especially for the syllable finals, are usually perceived as somewhat indistinct, it is due to the effect of neutralization. Specifically, the consonants are obviously shortened and reduced compare with that in corresponding normal syllables. In addition, it is most likely to become voicing; At the same time, the vowel quality in NT syllable is neutralized and sounds close to that of the shwa. According to Fig. 2, however, their original sound quality is still identified. Fig. 2 shows a comparison in the

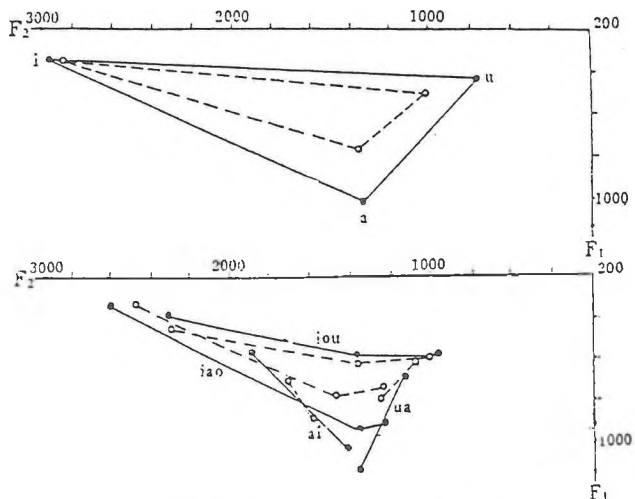


Fig 2. Comparison between NT and normal syllables in their acoustic vowel-chart

related acoustic vowel-chart between NT syllables and the normal ones. The top one represents the position of the monophthongs /i/, /a/, /u/, and the bottom one of diphthongs /iao/, /ua/, /iou/ and /iou/, here the solid line marks the vowel quality in normal syllables, and the dashed line marks that in NT syllables. As what can be observed from this Figure, the vowel quality of NT syllable in all cases are reduced just towards the target of shwa compare with the original qualities in normal syllables, but still keep in discriminable one another.

### III. CONCLUSION

The experimental results obtained from this investigation leads to following considerations:

- (a) As the hypothesis made at the beginning of this paper, the perceptual feature of neutral-tone syllables is caused by the combined effect of neutralization taken place in all the aspects of the syllable including intensity, pitch, duration and the sound quality, rather than any single factor;
- (b) As a special morpheme and differentiate from the common unstressed syllables, the neutral-tone syllables in Mandarin Chinese do have some relatively fixed patterns in acoustic realization, and it is mainly determined by tonal difference of the preceding syllables.

### REFERENCES

- [1] Dong, Shaowen. An Introduction to Phonetics. Cultural & Educational Press, 1956, pp. 80-82.
- [2] Lin, Maocan & Yan, Jingzhu. The acoustic characteristics of neutral-tone syllables in Standard Chinese. Fangyan, No.3, 1980, pp.166-178.
- [3] Lin, Tao. Experimental approach on neutral-tone syllables in Standard Chinese. The Collections of Linguistics papers, Vol.10, 1982, pp.16-37.
- [4] Cao, Jianfen. Analysis of neutral-tone syllables in Standard Chinese. Applied Acoustics, 5:4, 1986, pp.1-6.
- [5] Chao, Yuanren. A Grammar of Spoken Chinese, University of California Press, 1968.
- [6] Chao, Yuenren. On the Language, Commercial Press, 1980, pp.59-83.
- [7] Luo, Changpei & Wang, Yun. An Outline of General Phonetics, Commercial Press, 1981, pp. 133-135.
- [8] Wu, Zongji & Lin, Maocan ed. An Outline of Experimental Phonetics, Higher Educational Press, Beijing, 1989.
- [9] Yang, sun-an. Preliminary synthesis of neutral-tone syllable in Standard Chinese. PL-ARPR/1988, pp.81-91.
- [10] Cao, Jianfen. Temporal structure in bisyllabic words frame: An evidence for relational invariance and variability from Standard Chinese. The Proceedings of ICSLP92, Alberta, Canada, 1992.