Prosodic Encoding of Topic and Focus in Mandarin

Bei Wang[1] and Yi Xu[2,3]

[1]University of Potsdam, Germany
[2]University College London, UK, [3]Haskins Laboratories, USA
bwang@uni-potsdam.de, yi@phon.ucl.ac.uk

Abstract

In this study, we investigate whether and how focus and topic can be separately encoded in Mandarin. A total of 60 sentences with three lengths and five tone combinations were recorded in four topic-focus conditions: initial focus, new topic, implicit topic and given topic, by six speakers. The results of acoustic analysis show that new topic is encoded with a *raised* pitch range on the initial word. Focus, in contrast, is encoded with an *expanded* pitch range on the focused word and a suppressed pitch range on the subsequent words.

1. Introduction

Topic and focus are two important concepts of information structure in speech. They differ in their pragmatic functions and often in their syntactic forms as well [1: 206]. Hallidy[2] defines focus as one kind of emphasis, whereby the speakers marks out a part of a message block as that which they wish to be interpreted as informative. The topic of a sentence is the thing which the proposition expressed by the sentence is ABOUT [1: 118].

Frequently, however, focus and topic are not manifested with distinct syntactic forms [3]. The question then arises as to whether there are phonetic means by which to encode the functional contrasts otherwise expressed syntactically. Many studies have shown that intonation is used to realize information structure. It is well known that a focused word has higher F_0 and longer duration. Also, though much less well known, there is a sharp post-focus F_0 drop, and the F_0 of each subsequent word in the reminder sentence is lowered [4][5]. It is further suggested that focus is realized through a tri-zone pitch range control: expanding pitch range of focused component, suppressing pitch range of post-focus components, and leaving pitch range of pre-focus component neutral [6][7][8]. For topic, Féry [9] claimed that the topicalized constituent has a prominent rising accent on its accented syllable and it is separated from the rest of the sentence by a boundary tone, or even by duration.

A contrastive focus may occur in any position in a sentence [3]. When occurring in sentence initial position, it may conflict with topic in acoustic manifestations. In [5], it was found that, in a long sentence, an initial focus does not raise the initial F_0 because it is already quite high without it. Chen [10] proposed that there is an upper limit on the amount of pitch rising by a focus. These findings suggest that the amount that F_0 is raised by a new topic is positively related to sentence length, and beyond a certain length, the amount of pitch rising by a new topic may exceed that by an initial focus.

To make things more complicated, the topic/focus distinction is often confounded by an alleged functional division between given and new information. It has been said that for West Germanic languages like English and German new information is marked by a pitch accent, while given information is deaccented [11]. A set of perception experiments on German has suggested that the type of pitch accent plays a role in the marking of different degrees of givenness [12]. Horne [13] observed that the width of the F_0 register on the subject was considerably greater in the 'new' context than in the 'given' context of British English, but not American English. There are also suggestions that the distinction between given and new information is not an either-or dichotomy but rather a continuum [14].

What is yet lacking are investigations in which topic, focus and given/new are systematically controlled. The goal of the present study is thus to find out whether and how focus and topics with different degrees of newness are encoded acoustically in sentences that are syntactically identical.

2. Method

The procedure used in this experiment involved reading aloud of sentences under four topic-focus conditions: initial focus (IF), new topic (NT), implicit topic (IT) and given topic (GT).

2.1. Reading Materials

The basic stimuli consisted of 5 base sentences with different tone combination (HH, HL, LH, FR, and RF) in Mandarin. They all had simple SVO syntactic structure and consisted of only disyllabic content words. The first word was always the subject of the sentence. According to [1:132], subjects are unmarked topics and topic-comment articulation is unmarked pragmatic sentence articulation. Based on every base sentence, three sentences with the length of 6, 14 and 20 syllables were composed.

For every sentence, priming contexts were composed to elicit the four targeted topic-focus variations, as listed below. Initial focus was elicited by a rebuttal paradigm, as in (1). New topic, which typically occurs at the start of a conversational turn or a monolog, was elicited with no priming context before the target sentence. Implicit topic was elicited by a priming context that implies but does not explicitly mention the subject of the target sentence. Given topic was elicited by a priming context

that explicitly mentions the subject of the target sentence. For instance, for the base sentence "Wang1Ying1 can1guan1 che1jJian1" (WangYing visits the factory), where WangYing is the subject, the four conditions are as follows:

(1) **Initial focus (IF):** Bu2 shi4 Li3Xiao3. (It is not Lixiao.), Wang1Ying1 can1guan1 che1jJian1.
(2) **New topic (NT):** Wang1Ying1 can1guan1 che1jJian1.
(3) **Implicit topic (IT):** Da4jia1 dou1 you3 shi4qing yao4zuo4. (Everybody has something to do.) Wang1Ying1 can1guan1 che1jJian1
(4) **Given topic (GT):** Wang1Ying1 he2 Li3xiao3 dou1 you3 shi4qing yao4zuo4. (WangYing and LiXiao both have something to do.). Wang1Ying1 can1guan1 che1jJian1.

Sentence length was varied by adding modifiers before the object, leaving the subject and verb of the sentences the same. Below is an example of one such group of sentences.

- **6 syllables:** Wang1Ying1 can1guan1 che1jian1. (WangYing visits the factory.)
- **14 syllables:** Wang1Ying1 Can1Guan1 Shan1Xi1 Qing1Xiang1 Yi1Jie1 Jia1Gong1 Che1Jian1. (WangYing visits ShanXi QingXiang YiJie hardware factory)
- **20 syllables:** Wang1Ying1 Can1Guan1 Shan1Xi1 Qing1Xiang1 Yi1Jie1 Fei1Ying1 Gong1Si1 Ji1Xin1 Jia1Gong1 Che1Jian1. (WangYing visits ShanXi QingXiang YiJie FeiYing company crankshaft hardware factory.)

Altogether, there were a total of 5 (tone combination) × 3 (sentence length) × 4 (contexts) = 60 sentences.

2.2. Participants

Participating in the recording experiment were 4 female and 2 male university students, aged from 22 to 25. All were native speakers of Mandarin Chinese with no self-reported speech and hearing impairments.

2.3. Recording procedures

The participants were recorded individually in a quite room. Each wore a head-mounted microphone. All sentences were recorded directly into a computer and saved as wav files. During each trial, a sentence was shown on the computer screen, and the subject was asked to say it in a natural way at normal speech rate without pause. For each speaker, the reading material was repeated four times in random order. Only the last three repetitions were used for data analysis.

2.4. Acoustic Measurement

A Praat script was used to analysis the acoustic parameters of target sentences [15]. To measure F_0, the vocal cycles were first marked by Praat [16] and then hand checked for errors and finally smoothed by the script. Syllable boundaries were also labeled by hand. The highest pitch points of the High tone (H) and the Falling tone (F) were measured as the maximum F_0 of the word; and the lowest pitch point of the Low tone (L) and the Rising tone (R) as the minimum F_0.

3. Results

3.1. Pitch and duration of the first words

The maximum F_0 of the first words are shown in Figure 1, where the means are arranged according to sentence length and topic-focus condition. Each value in the figure represents the value of 5 sentences with 3 repetitions averaged across all six speakers. Cross speaker paired sample T tests were carried out between each pair of the topic-focus conditions. Figure 1 shows that the maximum F_0 of the first word is higher when it is focused than in the other three conditions.

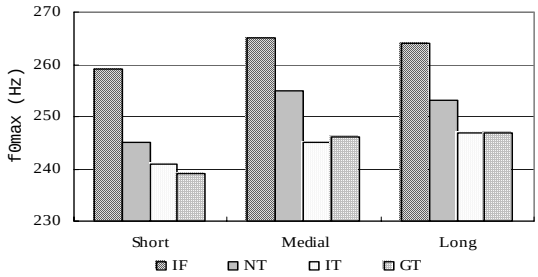

Figure 1. *Maximum F_0 of the first word.*

The results of the paired-sample T-test are presented in Table 1. (Here and subsequently, * stands for $p < 0.05$, ** for $p < 0.01$, and *** for $p < 0.001$.) The differences in maximum F0 between initial focus and the other three conditions are all significant. This is true for all sentence lengths. As for the three topic conditions, the only significant difference are seen between new topic and implicit topic of the medial sentences, and between new topic and given topic of the long sentences.

When the initial words are in focus, maximum F_0 is significantly higher in the medial and long sentences than in the short sentence (T(5)=-4.051* for short vs. medial sentences, and T(5)=-3.922* for short vs. long sentences), but the difference between medial and long sentences is not significant (T(5)=0.99). It is the same pattern with new topic (T(5)=-3.727* for short vs. medial sentences; T(5)=-2.997* for short vs. long sentences; and T(5)=0.682 for medial vs. long sentences).

Table 1. *Results of paired sample T-test on the maximum F_0 of the first word, value of T(5).*

	Short	Medial	Long
IF-NT	3.671*	3*	3.573*
IF-IT	5.832**	5.863**	5.699**
IF-GT	8.61***	4.742**	5.157**
NT-IT	1.244	2.734*	1.83
NT-GT	2.324	1.711	3.22*
IT-GT	1.807	-0.266	0.084

Minimum F0 values of the first words are presented in

Figure 2. As can be seen, the new topic condition always has the highest minimum F0. There is a significant difference in the paired sample T test between focus and new topic in the short (T(5)=-3.288*) and the medial sentences (T(5)=-3.11*). Unlike maximum F0, the differences in minimum F0 are significant between new topic and implicit topic (T(5)=3.56*) and between new topic and given topic (T(5)=4.11**) in the medial sentences. This also holds for long sentences (new topic vs. implicit topic: T(5)=2.496*), and new topic vs. given topic: T(5)=2.953*).

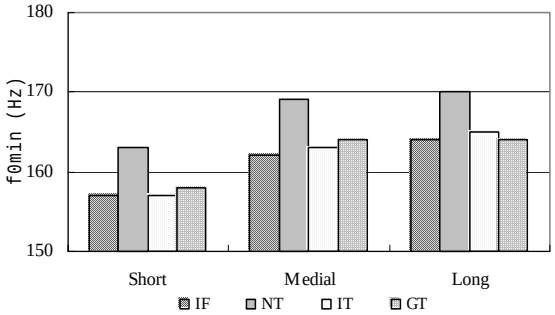

Figure 2. *Minimum F_0 of the first word.*

The duration values in Figure 3 show that when the initial word is focused, it is significantly longer than in the other three conditions. Also, initial words in implicit topic are significantly longer than in given topic of the short sentences (T(5)=3.172*) and the medial sentences (T(5)=2.958*), however, not significantly longer in the long sentences (T(5)=1.564). There is no significant difference in duration between new topic and implicit topic at any sentence length. And, only in the short sentences, is new topic significantly longer than given topic (T(5)=3.234*).

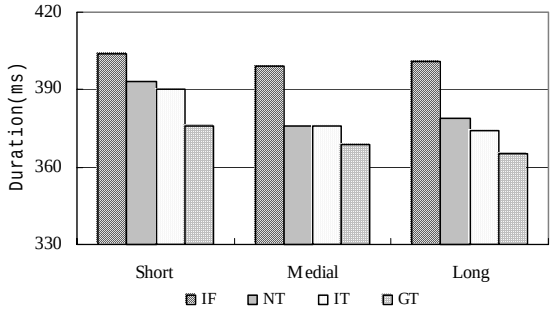

Figure 3. *Duration measures of the first word.*

In summary, in initial focus the first word has larger pitch range mainly due to raised maximum F0, and longer duration. There are no significant differences in maximum F0 among other three conditions. However, new topic has higher minimum F0 than given topic and implicit topic in the medial and the long sentences. The only difference between implicit topic and given topic is that the former is longer than the latter. But this is only true for the short and the medial sentences.

3.2. Pitch and duration of the second and subsequent words

The second words are verbs in all the sentences. The maximum F_0 values of the second words are presented in Figure 4. They are arranged according to sentence length and focus-topic conditions.

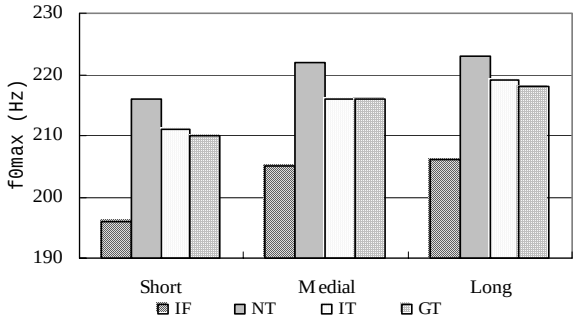

Figure 4. *The maximum F_0 of the second word.*

Paired sample T-tests reveal that when the first word is in focus, F_0 of the second word is significantly lower than in the other three conditions across all sentence lengths. Concerning sentence length, F_0 of the second word of short sentences is significantly lower than that of the medial (T(5)=-2.762*) and long (T(5)=3.942*) sentences. Thus when sentences are longer, the lowering effect is not as big as when they are shorter.

Among the other three conditions, the significant contrasts are: new topic vs. implicit topic in short (T(5)=2.609*) and medial sentences (T(5)=4.669*), and new topic vs. given topic in medial (T(5)=5.414*) and long sentences (T(5)=4.566*).

The similar analysis of minimum F_0 reveals that significant contrasts are only found in medial sentences, i.e., between new topic and implicit topic, between new topic and given topic, and between initial focus and new topic. This is in concert with the results of the first words.

As for duration of the second word, the only significant distinction is that new topic is longer than in the other three conditions. This means that initial focus does not have an effect on the duration of the following word.

In summary, the main finding about the second word is that a focused word lowers the pitch of the word after it, and the lowering is greater for the short sentences than for the medial and the long sentences. Moreover, the F_0 lowering is not accompanied by a shortening of the second word.

3.3. Temporal span of compressed pitch range after focus

To examine how long the pitch lowering effect of the initial focus lasts, we carried out paired sample T-tests on the F_0 from the third word to the tenth word between each pair of the four topic-focus conditions. Almost all the words in the initial focus condition are significantly lower than their counterparts in the other three conditions. This holds true even for the last words in the 20-syllalbe sentences. This indicates that the pitch lowering effect after focus persists through the whole post-focus region.

4. Discussion and Conclusion

In this study, we see that an initial focus raises the maximum F_0 and without raising the minimum F_0 of the first word. It also lowers the maximum F_0, but not the minimum f0, of all the subsequent words. In contrast, a new topic raised the minimum F_0 of the initial word without lowering the F_0 of the subsequent words. Focus seems to expand the pitch range of the focused word and compress the pitch range of the post-focus words. This is consistent with previous findings [4][6][7][8]. New topic seems to raise the pitch range of the initial word without lowering the pitch range of the subsequent words. The post-initial pitch range seems to fall gradually in a new topic sentence.

Our results here are not fully consistent with Cooper et.al.'s finding [4] that the initial pitch is relatively constant with or without initial focus. In our data, when the initial word is in focus, its pitch is higher than in the new topic condition. This result does not necessarily mean that initial focus always has higher pitch than new topic. It is likely that the experiment paradigm employed in the present study has not elicited the full effect of new topic. Larger magnitudes of pitch rising have been reported in sentences that start a new paragraph and a conversational turn [17][18][19]. Nevertheless, the new data, for the first time, have presented direct evidence on the critical difference between focus and topic in terms of their manner of encoding. The clear differences suggest that it is possible for the two functions to be perceptually recognized by listeners. Experimental verification of such parallel encoding will await future investigations.

In regard to the previously proposed given-new contrast, the best comparison is between the implicit and the given topic. In the former, the word is not directly mentioned but can be deduced from the context. In the latter, the word is explicitly mentioned in the context. Our results here show no significant difference in maximum F_0 or minimum F_0 between the two. The only difference is seen in duration: it is longer in implicit topic than in given topic. Even this difference is not significant in long sentences. And we do not yet know whether such durational difference is audible to listeners. Further studies are needed.

5. Acknowledgement

We are grateful to Caroline Féry for her constructive discussions and comments. We wish to thank Beijing TianLai Speech Data Company for the acoustic measurement of the speech data. The research was partly supported by the DFG project, SFB632: 'Information Structure'.

6. Reference

[1] Lambrecht, K., 1994. *Information structure and sentence form: Topic, focus and the mental representations of discourse referents*. Cambridge University Press.

[2] Halliday, M.A.K., 1967. Notes on transitivity and theme in English part II. *Journal of Linguistics* 3, 199-244.

[3] Bolinger, D. L., 1972. Accent is predictable (if you're a mind reader). *Language* 48, 633-644.

[4] Cooper, W.E.; Eady, S.J.; Mueller, P.R., 1985. Acoustical aspects of contrastive stress in question-answer contexts, *J. Acoust. Soc. Am* 77, 2142-2156.

[5] Eady, S.J.; Cooper, W.E., 1986. Speech intonation and focus location in matched statements and questions, *J. Acoust. Soc. Am.* 80, 402-415.

[6] Xu, Y., 1999. Effects of tone and focus on the formation and alignment of F_0 contours. *Journal of Phonetics* 27, 55-105.

[7] Xu, Y.; Xu, C. X., 2005. Phonetic realization of focus in English declarative intonation. *Journal of Phonetics* 33, 159-197.

[8] Xu, Y.; Xu, C. X.; Sun, X., 2004. On the Temporal Domain of Focus. *Proc. International Conference on Speech Prosody 2004*, Nara, Japan, 81-84.

[9] Féry, C., 2005. The prosody of topicalization. In. *ISAG volume*. Winkler, S; Schwabe, K.(eds.)Mouton de Gruyter.

[10] Chen, Y., 2003. *The Phonetics and Phonology of Contrastive Focus in Standard Chinese.* Ph.D. dissertation. State University of New York at Stony Brook. Stony Brook.

[11] Ladd, D. R.,1996. *Intonational phonology*, Cambridge Studies in Linguistics79. Cambridge: Cambridge University Press.

[12] Baumann, S.; Grice, M., 2004. Accenting accessible information, In *the Proceeding of Speech Prosody 2004*, Japan.

[13] Horne, M., 1990. Accentual patterning in 'New' vs 'Given' Subjects in English. *Working Papers*. Lund University, Dept. of Linguistics 36, 81-97.

[14] Chafe, W.L., 1976. Givenness, contrastiveness, definiteness, subject, topics, and point of view. In *Subject and topic*, C.N. Li (ed.). Academic Press, 25-53.

[15] Xu, Y., 2005. *http://www.phon.ucl.ac.uk/home/yi/downloads.html*

[16] Boersma, P.; Weenink, D., 2005. *http://www.praat.org/*

[17] Lehiste, I., 1975. The phonetic structure of paragraphs. In *Structure and process in speech perception*. Cohen A.; Nooteboom, S. E. G. (eds.). Springer-Verlag: New York**,** 195-206.

[18] Nakajima, S.; Allen, J. F., 1993. A study on prosody and discourse structure in cooperative dialogues. *Phonetica* 50, 197-210.

[19] Umeda, N., 1982. "F0 declination" is situation dependent. *Journal of Phonetics* 10, 279-290.