Prosodic Realization of Split Noun Phrases in Mandarin Chinese Compared in Topic and Focus Contexts

Bei Wang
Department of Linguistics, University of Potsdam, Germany

Abstract
The present study investigates the prosodic realization of split noun sentences in Chinese, like 'shu, wo mai le san ben.' (Book, I buy ASP three CLAS. 'I bought three books.'). The question and answer paradigm was used to induce sentences where the noun was either the topic or the focus. In the topic context, the question was of the form, 'I heard you bought books and pencils. Is that true?' In the focus context, it was of the form, 'What did you buy three?' Speech production results show that the prosodic realization of split noun sentences makes the noun a separate intonational phrase. The intonational difference between these two contexts for split noun sentences is not significant. The interesting finding is that the pause after the noun in the topic context is significantly shorter than in the focus context (189ms and 248ms on average, respectively), which shows that the base part is more closely related to the noun in the topic context than in the focus context. In both contexts, there is a pitch accent on the modifier in the base part. These findings give prosodic evidence for Pan's (2003) argument that a focus in the base is a cue to distinguish split sentences in a topic context, and whether the prosodic characteristics are sensitive to the focus. The second goal in this paper is to study to what extent the split-noun sentence is accepted by listeners and whether the prosodic characteristics are sensitive to the focus. The present study investigates the prosodic realization of split noun structures in Chinese using a speech production paradigm of matched questions and answers.

1. Introduction
In this paper, the prosodic realization of split noun structures in Chinese is studied from both a speech production and a speech perception approach. In Chinese a noun can be separated from its base position, as in (1b).

(1) a. Wo Mai le San Ben Shu. (canonical sentence)
   I buy ASP three CLAS book.
   b. Shu Wo Mai le San Ban. (split sentence)
   Book I buy ASP three CLAS.
   'I bought three books.'

Fanselow & Cavar (2002) studied split noun structure in many other languages and proposed two different kinds of split structure, that is, simple split and inverted split. In simple split, the order of the elements in the NP remains unchanged, while splits that have the reverse order are called inverted splits. Chinese allows inverted splits. Pan (2003) studied the asymmetry phenomena that noun movement is only allowed in relative clause (adjunct) + head noun cases but not in complement clause + head noun in Chinese.

He proposes that the asymmetry disappears if either a classifier is added to the relevant cases (see Pan’s example (2)), or the moved head noun is adjacent to its associated part at syntax or LF. He concludes that both of these situations are connected to the existence of focus in the relevant noun phrase. He claims that when the focus phrase that contains the empty head noun is adjacent to the topicalized NP, it establishes a relation with the topicalized NP and then the empty head noun is identified from the topic phrase to the focus phrase.

(2) a. Wo tingshuo -le [CP [NP Bill da -le Mary de] xiaoxi] (I hear ASP Bill hit ASP Mary De news)
   b. * Xiaoxii Wo tingshuo -le Bill da -le Mary de. (News I hear ASP Bill hit ASP Mary De.)

Fanselow & Cavar (2002) studied split noun structure in many other languages and proposed two different kinds of split structure, that is, simple split and inverted split. In simple split, the order of the elements in the NP remains unchanged, while splits that have the reverse order are called inverted splits. Chinese allows inverted splits. Pan (2003) studied the asymmetry phenomena that noun movement is only allowed in relative clause (adjunct) + head noun cases but not in complement clause + head noun in Chinese.

Féry (2005) suggests that the marked tonal pattern of split structures is an additional prosodic phrase and there might be two intonational patterns, that is, the fronted constituent is topicalized and the right part is generally focused, or, that the fronted part is focused and the rest is just deaccented. Concerning the information structure of the split constructions, fronting implies prominence, as a topic or a focus. Focus is for new information, the answer to a wh-question or a contrastive answer (Rooth 1985). Topic is used in its meaning of aboutness (Büring 1999). The topicalized element introduces what the remaining sentence is about, and it can be more or less clearly separated from the main clause (Féry, 2005). This paper aims to provide experimental evidence of the prosodic realization of split noun structures in Chinese using a speech production paradigm of matched questions and answers.

Pan’s study (2003) gave a detailed analysis that the split noun structure in Chinese is syntactically acceptable and that it is the result of head-movement. The second goal in this paper is to study to what extent the split-noun sentence is accepted by listeners and whether the prosodic characteristics are sensitive to the focus. The present study investigates the prosodic realization of split noun sentences in Chinese, like 'shu, wo mai le san ben.' (Book, I buy ASP three CLAS. 'I bought three books.'). The question and answer paradigm was used to induce sentences where the noun was either the topic or the focus. In the topic context, the question was of the form, 'I heard you bought books and pencils. Is that true?' In the focus context, it was of the form, 'What did you buy three?' Speech production results show that the prosodic realization of split noun sentences makes the noun a separate intonational phrase. The intonational difference between these two contexts for split noun sentences is not significant. The interesting finding is that the pause after the noun in the topic context is significantly shorter than in the focus context (189ms and 248ms on average, respectively), which shows that the base part is more closely related to the noun in the topic context than in the focus context. In both contexts, there is a pitch accent on the modifier in the base part. These findings give prosodic evidence for Pan’s (2003) argument that a focus in the base is a cue to distinguish split sentences in a topic context, and whether the prosodic characteristics are sensitive to the focus. The second goal in this paper is to study to what extent the split-noun sentence is accepted by listeners and whether the prosodic characteristics are sensitive to the focus. The present study investigates the prosodic realization of split noun structures in Chinese using a speech production paradigm of matched questions and answers.

Pan’s study (2003) gave a detailed analysis that the split noun structure in Chinese is syntactically acceptable and that it is the result of head-movement. The second goal in this paper is to study to what extent the split-noun sentence is accepted by listeners and whether the prosodic characteristics are sensitive to the focus. The present study investigates the prosodic realization of split noun sentences in Chinese using a speech production paradigm of matched questions and answers.

2. Speech production experiment
2.1 Methods
2.1.1 Participants
Seven native Mandarin Chinese speakers participated in the recording, two males and five females, 22-26
years old, from University of Potsdam who were not linguistic students.

2.1.2 Reading materials
In each question-answer set, contexts of focus and topic were constructed. In the topic context, two objects were mentioned in the question, like ‘I heard BAOXIN lost bookmarks and notebooks. Is that true?’, where it is clear that ‘bookmarks’ is the topic in an answer like ‘shu.qian, BAOXIN diu le san zhang.’ (‘bookmarks, BAOXIN lose ASP three CLAS.’), since ‘shu.qian’ was given in the context and it is what the answer was about. The focus context was introduced by wh-questions like ‘What did BAOXIN lose three?’ then in the answer, ‘shu.qian’ is the focus.

In each set, there were 4 sentences, T-C, F-C, T-S, F-S (see example in table 1).

Table 1. Example sentences of reading materials

<table>
<thead>
<tr>
<th>Topic question:</th>
<th>Canonical sentence:</th>
<th>Split sentence:</th>
</tr>
</thead>
<tbody>
<tr>
<td>I heard BAOXIN lost bookmarks and notebooks. Is that true?</td>
<td>BAOXIN diu le san zhang</td>
<td>shu1.qian1 (bookmarks)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BAOXIN lose ASP three CLAS</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(bookmarks)</td>
</tr>
<tr>
<td>Focus question:</td>
<td>F-C</td>
<td>F-S</td>
</tr>
<tr>
<td>What did BAOXIN lose three?</td>
<td>(focus question, canonical answer)</td>
<td>(focus question, split answer)</td>
</tr>
</tbody>
</table>

All the experimental sentences have simple SVO structure with 8-11 syllables where the subject is a person’s name with 2 syllables, the verb has 2 or 3 syllables, a modifier has 3 to 4 syllables and the noun is always a two-syllable word. To keep intonational variation simple, all the syllables in a sentence are the same tone except for functional words like de and le. Five basic sentences of each of the four lexical tones, H, R, LR, and F, are constructed. There are 5(set) ×4(tone:H, R, LR, and F) × 2(sentence type: canonical, split) × 2(context: topic, focus)=80 sentences in all. The questions were read and recorded by the author, who is a native speaker of Mandarin Chinese. 80 fillers were used in the experiment. The sequence of all sentences was randomized.

2.1.3 Recording
The questions were auditorily presented to the participants. They were asked to read the answer at normal speech as naturally as possible after hearing the question. All the sentences were printed out on papers. The speech was recorded to a computer directly with built-in 16 bit sound at a sampling rate of 22 kHz, using Cool-Edit software. Ten sentences were used as an exercise.

2.1.4 Speech data annotation
Praat was used for speech data annotation. The f0 was hand checked for every vocal cycle using Xu’s Praat script (1999), and the duration of vocal cycles was converted into f0 values automatically. The boundary of each word was manually labeled as well. Acoustic parameters of the subjects, the verbs, the modifiers and the nouns were calculated, like, the maximum f0, minimum f0, duration and the silence after the word. The high point of pitch is the maximum f0 of a word, while the minimum f0 is the low point of the pitch.

2.2 Results:
The high points of the pitch of four parts in the sentences with H, R and F tones are analyzed1. Results of canonical sentences are shown in figure 1 (T-C and F-C).

![Figure 1. F0 max of normal sentence](image)

As we can see in figure 1, when the noun is focused (F-C), it has a significantly higher maximum f0 than when it is in the context of topic (T-C) (F(1, 208)=11.46***). The maximum f0 of the rest of words is the same in topic and focus contexts (F(1, 208)=0.304, P=0.582, F(1,208)=0.001, P=0.98, and F(1,208)=0.623, P=0.431 of the subject, the verb and the modifier respectively). This is in consistence with Xu’s experiment (1999) that the prosodic realization of a focus is on-focus raising and post-focus lowering, and the f0 of the pre-focus words barely changes. In the topic context (T-C), the modifier is under focus, which creates a post-focus effect on the noun so that its pitch is lowered. This might be another reason why the noun has much lower pitch in topic context. Crucially, there is a clear difference in intonation for the canonical sentence between the topic and the focus contexts.

The high points of the pitch of split sentences are shown in figure 2 (T-S and F-S).

![Figure 2. f0 max of split sentence](image)

1 Because the focus on LR tone is realized by lowering the low point of the pitch (Xu,1999), sentences with LR tone are not included in the analysis of the high point of the pitch.

2 *** stands for P<0.001, ** is P<0.01, and * is P<0.05.
of the other words is slightly higher in topic contexts than in focus contexts, but the difference is not significant (F(1, 208)=0.786, P=0.376 of the subject, F(1.208)=0.438, P=0.509 of the verb and F(1, 208)=0.615, P=0.434 of the modifier). The important result is the obvious pitch accent on the modifier in the base part. That is to say, the intonational pattern of a split noun sentence is that the noun bears a pitch accent and the modifier in the base part is accentuated as well. The following analysis of the pause between the noun and the base part will show that actually they are two intonational phrases and that an accent in the base part is obligatory.

The results of the low points of the pitch show that there is no significant difference between topic and focus contexts for both normal sentence and split sentence. It gives more evidence to Wang’s (2002) arguments that the raising of the high point of the pitch reflects pitch accent in Chinese.

Figure 3 displays the duration of the noun, together with the length of silence after it and the duration of the base part of split sentences. Sentences of all four tones are included here.

![Figure 3. Duration of the split noun, silence after it and the base part](image)

It can be seen in figure 3 that the noun itself of the split sentence is separated as an intonational phrase. The interesting finding is that there is a significant difference between the silence in the topic context and that of the focus context (F(1, 278)=11.22***), but the difference of the duration of the noun (F(1,278)=0.125, P=0.724) and the duration of the rest of the sentence (F(1,278)=2.03, P=0.155) is not significant (F-S vs. T-S). It means that the base part is more closely related to the noun in topic context than in focus context. In topic context (T-S), the modifier (like ‘three’) is what the listener expects to know about the noun (like ‘bookmarker’). Without hearing the modifier, the addressee would consider that the answer is not complete yet. Meanwhile in focus contexts (F-S), an answer with only the noun is complete. The base part is simply a repeating of the question. In the case of focus context, the noun and the base part are in all likelihood two separate sentences.

To sum up, there is no significant difference of intonation of split sentences between topic and focus contexts. The significant difference is that the silence after the noun is much shorter in the topic context than in the focus context.

Experiment II studies the acceptance of split sentences and whether listeners are sensitive to the clue of silence length in focus and topic contexts.

3. Speech Perception Experiment

3.1 Method

3.1.1 Task

Question and answer pairs were played in sequence to participants. The question was either a focus question, or a topic question. The answer was either a canonical sentence or a split sentence recorded in the speech production experiment (T-C, F-C, T-S, and F-S). In each set, there were 2×4 question-and-answer pairs. Participants were asked to judge how good the answer matched the question considering intonation and sentence structure. A 10-point scale was used with 10 being the best. The whole experiment was run as following, which took about 1 hour. The sounds of one of female speakers in the production experiment were used in perception experiment. Seven basic sentences were selected for the formal experiment under the condition that the acoustic variation is just like the average results.

Part I. Eight basic sentences were used to make 64 question-and-answer pairs for exercise.

Part II. Seven basic sentences were used to make 56 question-and-answer pairs for the formal experiment.

Part III. Only split sentences in part II were used to make 28 question-and-answer pairs, and were tested twice.

3.1.2 Participants

Twenty-one participants from Beijing Normal University took part in the experiment. They were all native Mandarin Chinese speakers, 20-25 years old and did not have any hearing problems.

3.2 Results

The average value of the acceptability grades is shown in table 2.

<table>
<thead>
<tr>
<th></th>
<th>F-C</th>
<th>T-C</th>
<th>F-S</th>
<th>T-S</th>
</tr>
</thead>
<tbody>
<tr>
<td>Focus question</td>
<td>9.23</td>
<td>7.25</td>
<td>6.75</td>
<td>6.25</td>
</tr>
<tr>
<td>Topic question</td>
<td>6.19</td>
<td>5.74</td>
<td>5.19</td>
<td>4.99</td>
</tr>
</tbody>
</table>

In table 2, we can see that split sentences are acceptable, from 4.99 to 6.75 out of a 10-point scale. In Fanselow (2004), the acceptance of split sentences of German is 5 out of a 7-point scale, which equals 7.1 in a 10-point scale.

We can also see that the acceptance of sentences for the topic question is much lower than that for the focus question.

Moreover, when the question is a focus question, for canonical sentences the matched answer is significantly higher than unmatched answer (F(1,289)=67.875*** (F-C vs. T-C), but the difference is not significant for split sentences (F(1,289)=3.315, P=0.07) (F-S vs. T-S). Does this
mean that listeners are not aware of the difference in the length of silence in split sentences in topic and focus contexts.

To make sure that the judgment of split sentences was stable and to find out whether listeners can notice the difference of split sentence recorded in topic and focus contexts, all the split sentences were tested twice in part III. Results are shown in table 3.

Table 3. Average value of the judgment (Part III)

<table>
<thead>
<tr>
<th></th>
<th>III-1</th>
<th></th>
<th>III-2</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F-S</td>
<td>T-S</td>
<td>F-S</td>
<td>T-S</td>
</tr>
<tr>
<td>Focus question</td>
<td>6.83</td>
<td>5.91</td>
<td>6.77</td>
<td>5.88</td>
</tr>
<tr>
<td>Topic question</td>
<td>5.63</td>
<td>5.11</td>
<td>5.86</td>
<td>5.46</td>
</tr>
</tbody>
</table>

As can be seen in table 3, participants stay consistent in their judgments. The new result here is that when the question is a focus one, the matched answer is significantly better than the unmatched one (F(1, 278)=11.53***, P=0.001 and F(1, 278)=11.19***, P=0.001 for the first and the second judgment respectively). According to the results of the speech production experiment, there is no intonational difference of split noun sentences in topic and focus contexts. The only difference is the length of silence after the noun. Here we can see that listeners are aware of the length of the silence and prefer a longer pause after the noun in the context of a focus question.

4. Discussion

Split noun sentences are a marked structure in Chinese with a long pause after the noun. In general, they are used in specific contexts where the noun is topicalized or focused.

One point worth mentioning is that ‘BAOXIN diu le san zhang’ (BAOXIN lose ASP three CLAS [bookmarkers]) is a grammatical sentence in Chinese since Chinese allows noun ellipsis. A split noun sentence can thus be analyzed as being two sentences, which might be one reason that there is no difference in intonation for focus and for topic contexts. The pause between them indicates how closely the two parts are related. Actually, the noun and the base part of a split sentence are two intonational phrases and can even be interpreted as two sentences. The focus in the base part gives a clue that it is related to the topicalized noun.

The acceptance of a topic question is lower than the focus question. It could be that participants had problems understanding the question when it was designed to introduce a sentence with certain word as a topic. The relationship between focus and pitch accent is straightforward. The prosodic cues of topic are more complicated and subtle. Another reason that listeners do not like topic context might be that the answer is simply not a full answer. For a topic question like “I heard you bought books and pencils” a full answer should be like ‘Books, I bought three. Pencils, I bought two.’

5. Conclusions

To summarize, the results of these experiments show that the fronted constituent of a split sentence is topicalized and the right part is generally focused. There is no intonational difference of split sentences in topic versus focus contexts. The significant difference is that the silence after the noun is much shorter in the topic context than in focus context. Moreover, listeners prefer a longer pause after the noun when a split sentence is the answer to a focus question.

6. Acknowledgements

A big thank you for constructive suggestions and comments is to Caroline Féry. I am particularly grateful to Yi Xu for helpful discussions, and indebted to him for the Praat script. I would like to thank Lingling Su for the work of speech annotation, Torsten Roloff for the programs and Elizabeth Medvedovsky for English proof reading. All errors are mine.

This research is financed by the DFG project ‘Information structure (SFB 632)’ and ‘Morphosyntax and phonology of discontinuous nominal and prepositional constructions’.

7. References

[3]. Fanselow, G. and Frisch S., Stability and Plasticity of (gradient) acceptability judgments. Talk at the DFGs annual meeting in Mannheim, 2004
[4]. Féry, C., The prosodic and intonational structure of split DPs and PPs in a typological perspective, manuscript, 2005
[8]. Xu Y., Effects of tone and focus on the formation and alignment of f0 contours. Journal of Phonetics, Vol. 27, 55-105,1999