Multiple focus in Mandarin Chinese

Elena Kabagema-Bilan a, Beatriz López-Jiménez a, Hubert Truckenbrodt a,b,*

a Universität Tübingen, Seminar für Sprachwissenschaft, Wilhelmstr. 19, 72074 Tübingen, Germany
b Zentrum für Allgemeine Sprachwissenschaft (ZAS), Schützenstr. 18, 10117 Berlin, Germany

1. Background

1.1. Introduction

In this paper, we report the results of an experiment on the phonetic realization of multiple focus in Mandarin Chinese. The experiment builds on investigations of single focus in Mandarin Chinese by Xu (1999) and others, who show that a single (non-sentence-final) focus raises F0 in the focus and lowers and compresses F0 after the focus. Our question is whether these phonetic effects occur on both foci in a construction with two foci, or only on one of them. Our experiment stands in the context of the more general cross-linguistic question how phonetically measurable effects of focus are to be modeled, and how exactly they are to be formulated. Background to these issues is introduced in the following sections 1.2 and 1.3. Background to single focus in Mandarin Chinese is given in section 1.4. Previous results on multiple focus in English are reviewed in section 1.5. Section 2 presents the method of our experiment on multiple focus. The results are presented in section 3 and discussed in section 4. Section 5 sums up the main points of the paper.

1.2. Focus-marking by [F] and phonetic effects of focus: their place in the grammar

Phonetic studies of focus often describe their results in terms of phonetic consequences of the presence of focus. For example, focus leads to longer durations of the focused constituent and higher pitch peaks on the focused constituent (Eady...
et al., 1986; Rump and Collier, 1996 for English; Baumann et al., 2006; Baumann et al., 2007 for German; Xu, 1999; Yuan, 2004 for Mandarin Chinese).

A different picture emerges from the literature on the syntax–phonology interface (Truckenbrodt, 2007). Linguistic theorizing about effects of syntactic phrase structure on phonology has led to the widely held view that these are mediated by prosodic structure. As indicated in the first line of (1), the syntactic structure is mapped to a prosodic structure, which in turn can affect segmental phonology, though segmental phonology cannot be affected by syntax directly. This is called the indirect reference hypothesis, see, e.g., Inkelas (1989). Phonologists in turn like to believe that a model can be devised that relates their phonological representations (prosodic structure, segmental structure) to an idealized version of measurable phonetic realities. In intonation research, a suggestion for this relation has been made concrete in the pioneering work of Pierrehumbert (1980). This perspective on the relation between phonology and phonetics in intonation is influential in the large body of intonation literature that has grown out of Pierrehumbert’s work, see Gussenhoven (2004) and Ladd (2008).

(1) syntax -> prosodic structure -> segmental structure phonology

Since Jackendoff (1972), linguists assume that the feature [F] marks focus in the syntax. [F] is a syntactic feature, which has consequences in the semantics/pragmatics on the one hand, and in the phonology/phonetics on the other hand. In that suggestion, the feature [F] is part of the syntactic representation, as shown in (1). If these general assumptions, which are shared by many linguists, are put together without further ado, as in (1), then a direct link between focus [F] and the phonetic realities is not envisaged. Rather, phonetic focus effects would have to be indirect: The syntactic feature [F] would affect the prosodic structure, which in turn can affect the phonetic realities. For example, if focus leads to an increase in F0 height, the perspective in (1) allows only for an indirect way of modeling this: [F] must lead to some change in the prosodic structure, for example, an additional grid-mark of stress/prominence. In a second step, there would then need to be a general effect of such a prosodic property on the phonetics. For example, the presence of an additional grid-mark would generally lead to an increase in F0 height of the constituent that carries the grid-mark.

An alternative perspective might add a direct link between [F] in (1) and the box for phonetics there, i.e. allow a direct effect of [F] on the phonetics, such as a direct effect on the F0 height of a constituent. However, since the indirect reference hypothesis has been fruitful in the prosodic literature in other respects, we explore the possibility that it is correct for [F] as well.

1.3. The specific effects of [F] on the prosodic structure

For the effect of [F] on the prosodic structure, we take the suggestion of Jackendoff (1972) about the effect of a single focus on prosodic structure as the starting point: [F] attracts the strongest stress in its sentence. A refinement is argued for in Truckenbrodt (1995:Ch 4), see also Rooth (1992, 1996): The effect of stress attraction is limited to within the semantic domain of the focus (its scope), here called DF (domain of the focus). Truckenbrodt (1995:Ch 5) develops the perspective that stress-attraction to [F] within DF is the only direct effect of focus on the prosodic structure. Other consequences (such as the deletion of later beats of prominence in the focus domain or the insertion or deletion of prosodic boundaries) are derived indirectly from the resulting change in prominence relations under this suggestion. This perspective is pursued in interesting ways in Selkirk (2002, 2004), Sugahara (2005) and Büring (2009).

The current paper is concerned with multiple focus, as in the example (2a) from Schwarzschild (1999), where the verb would not require an accent if it were not focused. The suggestion of Schwarzschild (1999:170) is that each focus requires accent. This can also be seen in the pronominal examples in (2b), where the pronouns would not carry accent if they were not focused, cf. (2c).

(2) a. John cited Mary but he DISSED$_{p1}$ SUE$_{p2}$.
   b. I like YOU and HE likes HER.
   c. He LIKES her.

We transpose Schwarzschild’s suggestion minimally and work with the formulation that each focus requires a certain amount of stress, namely stress on the level of the phonological phrase. This is taken to be the amount of stress that is otherwise also required of each non-pronominal argument and adjunct of a verb. We assume that this level of stress entails the presence of an accent in English and German (see, e.g. Gussenhoven, 1992; Selkirk, 1995; Truckenbrodt, 2006).

If this is correct, it seems that the effect of [F] as attracting the strongest stress in its domain should be seen as a requirement not on each [F] (which requires only this lower amount of stress), but a requirement on the focus structure as a
whole, i.e. on DF: It requires that the strongest stress in it falls on some [F]-marked element. Among the F-marked elements in (2a), the decision which is the strongest is assumed to be made by a general preference for rightmost stress in the sentence, in favor of the rightmost focus. We assume, then, the prosodic representation in (3) for (2a). Stress, or more generally, an abstract notion of prominence, is here represented in terms of grid columns of different height. This representation of prominence may be derived by the constraints in (4) on prosodic effects of focus and the constraint in (5) on intonation phrase stress; see also Truckenbrodt (in press).

(3)  
\[x_{\text{intonation phrase level}} x_{\text{phonological phrase level}} \ldots \text{DF [but he \ DISSED}_{F1} \ SUE_{F2}}\]

(4)  
a. **Focus**: Each DF must have its strongest stress on some [F]-marked constituents inside of it.

b. **Stress-F**: Each [F]-marked constituent must carry stress at the level of the phonological phrase.

(5) The strongest stress in the intonation phrase is on the last phonological phrase level beat of stress.

A different assessment of the stress pattern of multiple foci is proposed by Selkirk (2005). Selkirk (2005) considers sentences like (6) in the context of the board game Clue, which defines a set of possible perpetrators, a set of possible locations, and a set of possible weapons. Players make suggestions or accusations. Selkirk suggests that, in doing so, they answer the implicit question ‘Who did it where with what?’ (6) is an example utterance of this kind. Selkirk observes that sentences of this kind are naturally produced with intonation phrase divisions and correspondingly intonation phrase stress on each of the foci separately. She suggests, accordingly, that each focus carries intonation phrase stress.

(6) \( \left( x_{\text{intonation phrase level}} x_{\text{phonological phrase level}} \right)_{IP} \)

I suggest that the crime was committed [in the lounge] [by Mr. Green] [with a wrench].

This is different for answers to multiple questions as in (8), where a natural rendition employs a lower level of stress on the non-final focused constituents.

(8) \( \left( x_{\text{intonation phrase level stress}} x_{\text{phonological phrase level stress}} \right)_{IP} \)

‘Hans read the newspaper to Maria, and Peter read a book to Claudia.’
The latter stress-pattern is indistinguishable, to the native intuition of the third author, from a stress-pattern in an all-new sentence as in (9).

(9) [What happened?]

\[\text{x } \text{IP intonation phrase level stress} \]
\[\text{x } \text{x } \text{x phonological phrase level stress} \]

[Hans hat Maria die Zeitung vorgelesen] \[F \]
'Hans read the newspaper to Maria.'

It therefore seems to us that Schwarzschild's formulation is correct for German as well. Focus requires accent (here: phonological phrase level stress). The arguments in (8) are independently accented (cf. (9)) and so the requirement on the foci is satisfied vacuously. We therefore retain, for the moment, the working hypotheses in (4) and (5). We think that the stress-pattern in (6) is correlated with additional interpretational effects, which suggest a minimally different focus structure. First, there is a sense in which each focus in (6) is a separate claim, and thus perhaps part of a separate assertion at some level of abstraction. Second, there is a sense in which each of the foci in (6) may also be seen as answering the same implicit question 'How was the crime committed?'. While we have no suggestion worked out in detail for (6), we see a plausible direction for representing the different focus structures as in (10) and (11). In (10) (for (8) or more generally examples of the kind brought up by Schwarzschild) the foci genuinely form some sort of interpretative unit that is jointly associated with the focus structure as a whole (here represented by DF). The alternatives in this case are alternatives to the focus tuple, for example <Hans, Maria, die Zeitung> vs. <Peter, Claudia, ein Buch>. In (11), by contrast, each focus is part of a focus structure of its own. Each such focus structure corresponds to a separate claim by the speaker. The constituent [in the lounge] represents a claim about the location. This claim is assessed against alternatives for the location, and independently of the claim about the perpetrator and the murder weapon.

(10) DF i [DFj [DFk [\dots [Hans]_F-i \[\dots [Maria]_F-i [die Zeitung]_F-i \[\dots ]]]]

(11) DF i [DFj [DFk [\dots [\text{in the lounge}]_F-i [by Mr. Green]_F-j [with a wrench]_F-k]]]

If this way of making the distinction goes in the right direction, we still require an element of the suggestion by Selkirk (2005), namely that focus not only attracts the highest stress of some domain (as in (4a)), but in fact directly requires stress at the level of the intonation phrase. This is formulated in (12a). The formulation here (in difference to Selkirk's original and similarly to (4a)) only requires intonation phrase level stress on one among a number of foci, so as not to assign intonation phrase stress on each focus in (8)/(10). In (6)/(11), however, the formulation in (12) requires that each DF leads to the assignment of one such intonation phrase stress on its F, with the correct consequences that each focus in this case receives intonation phrase stress. We therefore revise (4) as in (12), changing the a.-part to include a requirement for intonation phrase stress within DF.

(12) a. Focus: Each DF must carry stress at the level of the intonation phrase on some [F]-marked constituent.

b. Stress-F: Each [F]-marked constituent must carry stress at the level of the phonological phrase.

In the resulting account, each F requires phonological phrase level stress by (12b), and each DF requires intonation phrase level stress on some F in it, by (12a).

1.4. Single focus in Mandarin Chinese

We turn to the phonetic effects of a single focus in Mandarin Chinese. We build on the results concerning single focus of Xu (1999).

Mandarin Chinese has four lexical tones (1 = high, 2 = rising, 3 = low-fall(+rise), 4 = fall) and neutral tone. Xu (1999) investigates short sentences of Mandarin Chinese (a bisyllabic sentence subject + a monosyllabic verb + a bisyllabic object) with lexical tones permuted systematically in all positions; his study compares unfocused recordings with recordings that have narrow focus in different positions. Xu documents two phonetic main effects of focus in non-sentence-final position, which are observed for all lexical tones. We illustrate the two effects here with the tonally simplest case—sentences with five H tones (tones 1) as in Fig. 1.

The no-focus condition (thin solid line) is approximately horizontal. It serves as the baseline for the evaluation of focus. Relative to this baseline, the two phonetic consequences of focus in non-sentence-final position (on the subject or on the
verb) are: (i) raising of the element in focus and (ii) lowering and compression of material following the focus. Thus, subject focus (thick dashed line) leads to raising of the phonetic height of the H tones in the subject, and to lowering and compressing the phonetic height of the H tones of the following verb and object. Verb focus (thick solid line) leads to the same effects: The height of the H tones of the verb are raised by verb focus, and the height of the H tones of the following object are lowered and compressed. Verb focus shows in addition that material preceding the focus is not phonetically affected: The height of the H tones of the subject are not affected by verb focus. (Focus in final position (thin dashed line) differs: there is no later material of which the range could be lowered and compressed; at the same time, the raising in the focus itself is also weaker.) The effects of non-sentence-final focus in (i) and (ii) are found for all tones in the study of Xu.

Similar findings concerning the effect of focus on the overall pitch contour were reported by Gårding (1987) and Shih (1988, 2004) and confirmed by the study of Yuan (2004). Evidence for similar pitch range adjustments in question intonation are reported in Liu and Xu (2005). Aside from the expanded pitch range at the focus, Shih (1988) notices that the prominent syllables also have longer duration and higher intensity. The measurements of Xu (1999) also show a significant increase in duration under focus for the subject, for the verb, and for the object.

Where in the grammar are these phonetic effects of focus in Mandarin Chinese to be located?

If the feature \([F]\) can affect the phonetics directly, its phonetic consequences may be modeled by postulating that non-sentence-final \([F]\) raises the \(F_0\) in the focused element, and lowers and compresses the \(F_0\) range that follows \([F]\). Increased duration in the focus would likewise be a direct effect of the focus-feature \([F]\).

On the perspective in (1), a different way of modeling this effect would need to be pursued. This path follows, for Mandarin Chinese, the terminology of Shih (1988, 2004), Kochanski and Shih (2003) and Jin (1996) who discuss the phonetic correlates of focus in terms of prominence (Shih, 1988), stress (Jin, 1996) and prosodic strength (Kochanski and Shih, 2003; Shih, 2004). An analysis in terms of focus-induced metrical prominence is also discussed as a possible analysis for Mandarin Chinese in Chen and Gussenhoven (2008). For the purpose of this paper, we take as a starting point that (12) would be applicable in Mandarin Chinese as well and requires intonation phrase level prominence on a focused element. (13) shows a schematic representation of this kind. The raised \(F_0\), the following compression, and the change in duration, would then need to be phonetic consequences of the intonation phrase level prominence (rather than of the focus feature \([F]\)).

\[
(13) \quad x \quad \text{intonation phrase level prominence} \\
\quad \text{[DF subject [verb]F object]}
\]

Since no comparable effect is evident when narrow focus is not involved, one would be led to maintain that intonation phrase level prominence is not assigned if no narrow focus is involved, as in (14).

\[
(14) \quad \text{no intonation phrase level prominence assigned} \\
\quad \text{[subject verb object]}
\]

This would be a natural move in the context of the recent arguments in Selkirk (2005) and Katz and Selkirk (2009) that prosodic representations without intonation phrase level prominence are the default in English. In these suggestions, the last among a number of accents is not strengthened at the intonation phrase level, in difference to various earlier suggestions.

Our experiment on multiple foci in Mandarin Chinese is set in the context of these deliberations. Consider the structure in (15), which we tested in our experiment. There are two non-sentence-final foci, one on the subject, the other on the modifier of the object. (15) also shows the expectations of a prosodic account that employs (12). By (12), a single beat of intonation phrase prominence would be assigned (to the rightmost of the two foci), in parallel to (3) and (8). This beat of intonation phrase prominence in turn triggers \(F_0\) raising in the prominent position and later lowering and compression. No raising and lowering is expected for the first focus, since only a single beat of intonation phrase prominence is assigned by (12).
(15) Expectation of the prosodic account (12): one intonation phrase prominence for complex focus; phonetic raising and following lowering/compression for the intonation phrase prominence

\[ \text{[subject]}_F \text{ verb } [\text{[modifier]}_F \text{ noun}]_{\text{object}} \]

intonation phrase level prominence

A direct account of the phonetic consequences of [F] makes markedly different predictions, as shown in (16). If the feature [F] directly triggers phonetic F0-raising and following lowering/compression, we expect these phonetic effects to occur with each F-feature.

(16) Expectation of a direct phonetic account:
phonetic raising and following lowering/compression for each [F]-marked constituent

\[ \text{[subject]}_F \text{ verb } [\text{[modifier]}_F \text{ noun}]_{\text{object}} \]

Notice also that a prosodic account in which a beat of intonation phrase prominence is assigned to each focus in such a case would make the same prediction as the direct phonetic account in (16). The intonation phrase prominence that both the first focus and the second focus would have in such an account would lead to raising and following compression on both foci.

1.5. Earlier phonetic studies on multiple focus

Eady et al. (1986) and Rump and Collier (1996) have investigated the phonetic correlates of multiple foci in English and in Dutch. Eady et al. (1986) investigated sentences of the form [subject verb direct-object prepositional-object] in English, such as Jeff gave the ball to the cat. Double focus on subject and prepositional object was compared to a focus-neutral condition and to single focus on either subject or prepositional phrase. They observed an F0 increase on each focus, including both foci in the multiple focus condition. (Duration was also measured and found to be increased on each focus; however, this occurred in the context of other unexpected duration results: the final prepositional phrase was longer when the subject was focused; also, in another experiment in the same paper, VP focus led to longer duration not only in the VP but also in the preceding subject of the sentence.)

Rump and Collier (1996) investigated Dutch using the sentence A’manda gaat naar ‘Malta (A’manda is going to ‘Malta). By default, this sentence contains two accentual peaks, one on the sentence subject, one on the object. The subjects of the experiment were asked to adjust the height of one peak, given the height of the other peak. The height-adjustment was to be made in such a way that an optimal-sounding realization of a given focus-condition was achieved: focus-neutral, single focus on the subject, single focus on the object, or multiple focus on the subject and on the object. These were distinguished by context questions. The authors gave an overview of their results in the form of averaged (prototypical) pitch contours obtained, which are reproduced in Fig. 2.

Here, as in the study of Eady et al. (1986), double focus led to higher F0 of both foci in comparison to neutral focus. There are a number of possible reasons for this: (a) As in (16) above, each focus might directly trigger F0-raising. (b) By a constraint like Stress-F in (12b), the prominence on both foci might be strengthened relative to what it would be without focus; the larger prominence might lead to greater F0-height. (c) Bolinger (1986) suggested that a greater amount of speaker involvement raises F0 height, see also Hirschberg and Ward (1992). It is possible that focused utterances are felt to be more natural with a greater amount of speaker involvement than unfocused utterances, so that higher peaks may seem appropriate for double focus than under neutral focus.

Mandarin Chinese is interesting because it might allow us to distinguish among such possible sources of raising. As was seen in Fig. 1, the baseline of a focus-neutral rendition of an all-H-toned utterance is a near-horizontal phonetic realization in Mandarin Chinese. The height of this baseline is preserved in the syllables preceding the focus in late-focus conditions. In these cases, then, overall-raising of the phonetic height because of increased speaker involvement can be expected to also raise the height of this baseline to the left of a focus. Furthermore, the absence of peaks in the focus-neutral condition in Mandarin all-H (Tone 1) sentences suggests that lower levels of prominence, if they are assigned, do not affect the phonetic height of the tones. From a prosodic perspective, it would seem that only intonation phrase prominence has a clear effect on F0-height. Given an account along the lines of (12), we may therefore hope that the predicted asymmetry of prominence between non-final and final foci still emerges in an asymmetry of the presence of F0-effects. The prominence assigned to the first focus may still be below the threshold of prominence that leads to raising of the prominent element (and later lowering/compression), while the greater amount of intonation phrase prominence on the second focus may be strong enough to lead to F0 raising and later lowering/compression.

We now turn to our experiment on multiple focus in Mandarin Chinese.
2. Materials and methods

2.1. Stimuli

The four Chinese SVO sentences in (17) were designed for the experiment. All four sentences contain a disyllabic nominal subject and a monosyllabic verb. Sentences 1 and 2 contain an object that consists of a monosyllabic noun, preceded by a disyllabic nominal modifier. In sentences 3 and 4, both the object noun and its modifier consist of two syllables and were connected by the grammatical particle ‘de’.

All stimulus sentences were composed of syllables carrying high level tone (tone 1), except for the grammatical particle ‘de’ which has neutral tone. In order to avoid the disturbance and interruption of the continuity of F0 contours, an effort was made to use syllables with a sonorant initial consonant as much as possible, especially in the words that were to be focused.

(17) Stimulus sentences

<table>
<thead>
<tr>
<th>Sentence</th>
<th>Subject</th>
<th>Verb</th>
<th>Modifier</th>
<th>Gram. particle</th>
<th>Object noun</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sentence 1</td>
<td>māomī 'kitty'</td>
<td>tōu 'steal'</td>
<td>wūyā 'raven'</td>
<td></td>
<td>wŏ 'nest'</td>
</tr>
<tr>
<td>Sentence 2</td>
<td>gūmā 'aunt'</td>
<td>shāo 'cook'</td>
<td>dōngguā 'winter melon'</td>
<td></td>
<td>tāng 'soup'</td>
</tr>
<tr>
<td>Sentence 3</td>
<td>Zhāng Wēi 'Zhang Wei' (name)</td>
<td>lìn 'carry'</td>
<td>Wāng Yōu 'Wang You' (name)</td>
<td>de</td>
<td>shū-bāo 'school bag'</td>
</tr>
<tr>
<td>Sentence 4</td>
<td>Sūn Yīn 'Sun Yin' (name)</td>
<td>mō 'stroke'</td>
<td>Pān Ān 'Pan An' (name)</td>
<td>de</td>
<td>māomī 'kitty'</td>
</tr>
</tbody>
</table>

**Focus 1**

**Focus 2**

Fig. 2. Averaged adjustments (prototypical contours) reported by Rump and Collier (1996:9).
As indicated at the bottom of (17), focus is tested in the positions of the subject (first word) and the modifier of the object (third word, following subject and verb). The neutral tone syllable ‘de’ occurs in the syllables following the second focus position in sentences 3 and 4. The two different sentence lengths were not intended as contrasting conditions. Rather, we included the longer sentences 3 and 4 so as to have material with more than one syllable following the focus in the object, and we included the shorter sentences 1 and 2 so as to have material without the neutral tone on ‘de’.

Results of five different focus conditions in (18) are reported here. They are illustrated in (19).  

(18)  
1. Broad focus, where none of the words is emphasized (Broad).  
2. Multiple question-induced focus, where the subject and the modifier of the object together provide the information asked for by a wh-question (F-SU-M-q).  
3. Multiple contrast-induced focus, where the subject and the object modifier correct the content of a preceding yes/no-question (F-SU-M-c).  
4. Question-induced focus only on the subject (F-SU).  
5. Question-induced focus only on the modifier of the object (F-M).  

(19) Sentence 1 with its five different focus contexts

<table>
<thead>
<tr>
<th>Focus Condition</th>
<th>Question</th>
<th>Answer</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Broad</td>
<td>Yōu shénme xīn xiāoxi?</td>
<td>‘What’s new?’</td>
</tr>
<tr>
<td>2. F-SU-M-q</td>
<td>Shéi tōu shéi de wō?</td>
<td>‘Who steals whose nest?’</td>
</tr>
<tr>
<td>3. F-SU-M-c</td>
<td>Lāoyīng tōu xǐqùe wō ma?</td>
<td>‘Does an eagle steal a magpie’s nest?’</td>
</tr>
<tr>
<td>4. F-SU</td>
<td>Shéi tōu wūyā wō?</td>
<td>‘Who steals a raven’s nest?’</td>
</tr>
<tr>
<td>5. F-M</td>
<td>Māomí tōu shéi de wō?</td>
<td>‘Whose nest does a kitty steal?’</td>
</tr>
</tbody>
</table>

Target sentence for 1.–5. māomí tōu wūyā wō | ‘A kitty steals a raven’s nest’  

The contexts for sentences 2, 3, and 4 are shown in Appendix A.

2.2. Subjects

Five speakers of Mandarin Chinese, three women (LL, YL, YX) and two men (YZ, WS) were recorded for this experiment. Three speakers (LL, YX and WS) were born and raised in Beijing. One speaker (YL) came from Harbin, a city in Heilongjiang province (North of China), which belongs to the area of Northern Mandarin. The speaker YZ was born in YueYang city in the northern part of Hunan province where he lived until the age of 12 but spent another 12 years in Beijing before coming to Germany. All of the speakers were or still are students at Tübingen University, Germany. None of the speakers lived in Germany for more than five years (ranging from four months to five years). The age of the speakers ranges from 24 to 39 years.

2.3. Recordings

Recordings were conducted in a quiet room of the department of General Linguistics at the University of Tübingen. The subject was seated in front of the computer screen. The microphone was placed about 3 in. (7 cm) in front of the subject’s mouth. The target sentences together with the questions were printed in Chinese and displayed on a computer screen, one question-answer sequence at a time, in random order. Subjects were instructed to take their time to read silently the question-answer sequence in order to understand the connection between the two and to get prepared to read the answer appropriately. When the subject was ready, he/she had to press the ‘forward’ button to get the question played and to read aloud the target sentence as an answer. A native Mandarin Chinese speaker recorded the questions prior to the experiment, which were then connected to the PowerPoint presentation. The experiment was preceded by a short practice session consisting of three question-answer pairs resembling those used in the experiment. The speakers were controlling the pace of the experiment themselves and were instructed to repeat the recording whenever they felt they made a mistake or produced an unnaturally or improperly read answer. With the help of an external sound card Edirol UA-25, the recordings were saved directly into the computer with the Praat Recorder at a sampling rate of 44100 Hz.

---

1 A sixth condition in which the non-focused elements are contextually given, but the focused elements were not to be contrasted, was recorded as well, though with inconclusive results. At the advice of a reviewer, it is not included in this paper.

2 The northern part of Hunan province belongs to the area of South-western Mandarin.
Each of the four sentences are evaluated in five focus conditions. Each sentence in each focus condition was presented and recorded twice by each speaker. This amounts to 200 token sentences (4 sentences × 5 conditions × 2 repetitions × 5 speakers). Of these, five sentences from speaker YZ had to be excluded from further analysis because the second syllable of the word Zhāng Wēī (personal name), was pronounced with the falling tone instead of high level. Thus, a total of 195 sentences were taken for the analyses.

2.4. Measurements

Syllable boundaries were manually inserted with the help of spectrograms in Praat. These labels provided the basis for duration measurements. The middle point of the vocalic part of each syllable was labeled as well, and F₀ values were measured at these points. In a number of utterances, the F₀ value of the sentence final syllables (the last two to three syllables) in postfocal position could not be measured due to creaky voice.

The F₀-measurements were normalized for speaker-specific tonal height. For each speaker S, the average height of the first syllable in unfocused utterances, call it H₁ₛ, is first computed. The normalized values of this speaker are obtained by dividing each values of this speaker through the value of H₁ₛ. This procedure was chosen in the absence of a systematic low point in the data. It is felt to be sufficient for pooling of data across speakers and extracting across-speaker-tendencies from the pooled data.

3. Results

3.1. F₀ results

Fig. 3 shows the pooled normalized results of the five speakers, separated for sentence 1 (top left), sentence 2 (top right), and sentences 3 and 4 (bottom left). The sentences are numbered as in (17) and in Appendix A. Sentences 1 and 2 are plotted separately because they show a difference that we had not anticipated. We return to this below. The plotting points are the pooled averaged measurements at the beginning and end of the utterance, as well as the middle of the vocalic part of each syllable. Each focus condition is represented by 10 tokens for sentence 1, 10 tokens for sentence 2, and close to 20 tokens for sentences 3 and 4 together. (Close to 20 because of the five excluded tokens mentioned above; in detail: Broad: 19, F-SU-M-q: 19, F-SU-M-c: 18, F-SU: 19, F-M: 20).

The ‘baseline’ for comparison with the other conditions is broad focus, plotted by a thick solid line. This line is more or less horizontal in sentences 1 and 2. In sentences 3 and 4 it shows a lower point at “N6”; due to the neutral tone in this position.
The effect of single focus can be assessed in the conditions F-SU and F-M, both in comparison with broad focus. The overall effect of single focus in our data is comparable to the effect of focus in Xu (1999): F-0 in the focus is raised (relative to broad focus) and F0 following the focus is lowered or compressed (again relative to broad focus).

The conditions of particular interest are then those of multiple focus, either in response to a wh-question or by way of contrast (F-SU-M-q and F-SU-M-c). Notice first that visual inspection of the plots suggests that these two cases of multiple focus are not systematically distinct from each other. The plots further suggest that multiple focus on subject and modifier is not distinguished in the intonation from single focus on the modifier (F-M). Thus the plots show, in multiple focus of subject and modifier, some raising of the modifier and clear lowering or compression of what follows the modifier. However, the plots show little evidence of raising of the subject or lowering/compression of the verb, in the cases of multiple focus on subject and modifier. This is in contrast to raising of the subject and subsequent lowering where the subject is the only focus of the sentence (F-SU). Notice also that the amount of lowering/compression seems to be comparable across conditions, in a comparison of the sentence-final syllables: There are no systematic distinctions in final height between single and multiple focus conditions, or between early and late single focus conditions. Some of the preceding aspects are further assessed statistically below.

We turn to the difference between sentence 1 and sentence 2. It seems that the ‘normal’ case of raising in the focus is that the last (second) syllable of the focused constituent is raised most, and that the lowering/compression sets in after that syllable. We find this, in all three panels of Fig. 3, in the bisyllabic subject of the condition F-SU (plotting points H1 and H2 are the subject). We also find this for focus on the bisyllabic object modifier in sentence 1 and in sentences 3 and 4 in the condition F-M and in the multiple focus conditions. Here the plotting points H4 and H5 are the focused object modifier. Of these, H5 is raised in focus in sentences 1, 3, and 4. However, in sentence 2, H4 rather than H5 is raised in focus, and the lowering/compression already sets in inside of the focus in the position of H5.

We come to further statistical evaluations of the values that underlie the plots in Fig. 3. These statistical evaluations pool the data of sentences 1–4 at the following points:

- H2 of sentences 1–4: assessment of raising due to focus on the subject.
- H3 of sentences 1–4: assessment of lowering after a subject focus.
- H5 of sentences 1, 3, and 4, pooled with H4 of sentence 2 (see above): assessment of raising due to focus on the object modifier.
- H6 of sentences 1 and 2, pooled with H8 of sentences 3 and 4: assessment of lowering after focus.

For each of these points, an ANOVAs is computed with focus condition as fixed factor and normalized F-0 as dependent factor. The results are shown in Table 1, together with the significant distinctions of a post-hoc comparison.

The results of Table 1 can be summarized as follows.

First, for single focus, both the raising and the lowering/compression aspect are significant: Raising in the focused subject: F-SU is significantly higher at H2 than Broad focus, and than other conditions with later focus. Lowering/compression after the subject: F-SU is significantly lower at H3, H5/4, and H6/8 than Broad focus, and than other conditions with later focus. Raising in the focused object modifier: F-M is significantly higher at H5/4 than Broad focus. Lowering/compression after the focused object modifier: F-M is significantly lower than Broad focus in H6/8.

Second, in and around the subject position, single subject focus differs significantly from multiple focus on the subject and on the object modifier: single subject focus is significantly higher in the focused subject (H2) than focus on the subject and on the object modifier. Further, the verb (H3) is lowered significantly more after a single subject focus than in the cases of multiple focus on subject and modifier. At the same time, the multiple focus conditions are not significantly distinct from broad focus in these positions H2 and H3. Thus, the absence of phonetic focus effects with the first of two foci is statistically significant in the comparison with a single focus in the same position.

Third, multiple focus shows a significant effect of lowering/compressing after the second focus: The multiple focus conditions are significantly lower than broad focus in position H6/8. At the same time, the amount of lowering/compression

<table>
<thead>
<tr>
<th>H2 (subject)</th>
<th>H3 (verb)</th>
<th>H5/H4 (object modifier)</th>
<th>H6/H8 (last syll.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>df, F</td>
<td>df = 4; F = 14,1</td>
<td>df = 4; F = 36,9</td>
<td>df = 4; F = 89,9</td>
</tr>
<tr>
<td>Sig</td>
<td>p &lt; 0.001</td>
<td>p &lt; 0.001</td>
<td>p &lt; 0.001</td>
</tr>
<tr>
<td>Post hoc bearing on single F</td>
<td>F-SU &gt; Broad</td>
<td>Broad &gt; F-SU</td>
<td>F-M &gt; Broad</td>
</tr>
<tr>
<td></td>
<td>F-SU &gt; F-M</td>
<td>F-M &gt; F-SU</td>
<td>F-M &gt; F-SU</td>
</tr>
<tr>
<td>Post hoc bearing on two F</td>
<td>F-SU &gt; F-SU-M-q</td>
<td>F-SU-M-q &gt; F-SU</td>
<td>(F-SU-M-q &gt; Broad; p = .054)</td>
</tr>
<tr>
<td></td>
<td>F-SU &gt; F-SU-M-c</td>
<td>F-SU-M-c &gt; F-SU</td>
<td>Broad &gt; F-SU-M-q</td>
</tr>
<tr>
<td>Other sign. dist.</td>
<td>None</td>
<td>None</td>
<td>None</td>
</tr>
</tbody>
</table>

Table 1: Significance of distinctions due to focus conditions in the normalized data in the subject, the verb, the object modifier and the last syllable, assessed in the respective positions H2, H3, H5/4 and H6/8.
after the last focus is not significantly distinct between single focus cases and multiple focus cases (or between early and late focus cases). This confirms a similar visual impression pointed out in regard to Fig. 3 above.

Fourth, the results are less than conclusive in regard to whether multiple focus shows raising in the position of the last of the two foci: At a significance level of 0.05, the multiple focus conditions, assessed in the object modifier, are neither significantly distinct from broad focus nor from single focus on the object modifier. The plots in Fig. 3 do suggest some raising of the second of two foci on average, and the post hoc tests show a tendency in the same direction insofar condition F-SU-M-q shows higher F0 than Broad focus in the object modifier with a probability that approaches significance ($p = .054$). Thus, there is no strong evidence in support of raising of the second of two foci, though the data do not exclude this either.

In sum, in our data, multiple focus in two positions shows the same overall contour as single focus on only the second of these two positions. There is no raising and subsequent lowering/compression of the first of the two foci, while there is (less clearly) raising and (very clearly) subsequent lowering/compression accompanying the second of the two foci. This result obtains for multiple focus that is elicited with multiple wh questions, and it obtains for multiple focus that is elicited with multiple contrast in the context.

3.2. Duration results

Fig. 4 shows the averaged results of the duration measurements. The overall V-shape reflects the different inherent word lengths of the four constituents and possibly final lengthening. Of interest is the comparison among the conditions within each of the first three constituents, SU, V and Mod. This comparison is assessed with a separate ANOVA (mixed models) for each constituent with the dependent variable duration, the fixed factor focus condition, and the random factors speaker and sentence. The results are shown in Table 2.

In the position Mod the results are as expected if the last focus is lengthened: Where the modifier carries the only focus, it is longer than modifiers that are not focused. Further, where subject and modifier are focused, the modifier is also longer than in conditions in which the modifier is not focused. This supports the results of Xu (1999) that focus leads to an increase in duration in Mandarin Chinese.

The two conditions with multiple foci also show significantly increased relative length in the positions of the first of the two foci. However, this occurs in the context of two unexpected further results. For one thing, when the subject alone is focused, it is not significantly longer than the subject in the broad focus condition. For another, in the multiple focus conditions, the verb is significantly lengthened, along with the lengthening of the subject and of the modifier. We are not sure how to interpret this. The lengthening of the verb leads us to be cautious not to infer focus-induced lengthening of the first of two foci, since it would not be local. It seems possible that an extra-grammatical effect is at play. The speakers might have tended to slow down in the multiple focus utterances or parts thereof. We therefore will not draw any conclusions from these results here.

<table>
<thead>
<tr>
<th></th>
<th>SU</th>
<th>V</th>
<th>Mod</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single F vs. Broad</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>F-M &gt; Broad</td>
<td>F-M &gt; F-SU</td>
<td></td>
</tr>
<tr>
<td>Multiple focus</td>
<td>F-SU-M-c &gt; Broad</td>
<td>F-SU-M-q &gt; Broad</td>
<td></td>
</tr>
<tr>
<td></td>
<td>F-SU-M-c &gt; F-M</td>
<td>F-SU-M-q &gt; F-M</td>
<td></td>
</tr>
<tr>
<td></td>
<td>F-SU-M-q &gt; F-M</td>
<td>F-SU-M-q &gt; F-SU</td>
<td></td>
</tr>
<tr>
<td>Other sig. dist.</td>
<td>All other conds. &gt; F-M</td>
<td>None</td>
<td></td>
</tr>
</tbody>
</table>
The results also show unexpectedly short values for subject length in the condition of modifier focus. We leave this as an unexplained quirk in our duration data here. Xu (1999) reports that positions preceding a narrow focus are not shortened, and our study is not extensive enough to challenge that result.

4. Discussion

4.1. What accounts are compatible with the findings

While the duration results are not conclusive, the F0 results suggest that multiple focus in Mandarin Chinese shows phonetic focus effects for the second of two foci, but not for the first of the two foci.

The first of two foci on the sentence subject is statistically not distinct from the subject in the broad focus condition. The following verb (where lowering might occur) is also non-distinct in F0 in the two conditions. At the same time, the sentence subject has significantly higher F0 when it alone is focused than when it is the first of two foci. Similarly, the following verb is significantly lower when the preceding subject constitutes the only focus of the sentence than when the preceding subject is the first of two foci.

The second one of two foci, on the other hand, is followed by dramatic lowering that is not distinguished from lowering that follows a single focus, and significantly different from the broad focus condition. Raising of the second one of two foci is not conclusively shown by our data: the second one of two foci is not distinguished in F0 height from either the broad focus condition or from the condition that has a single focus in the same position. However, some raising is present numerically, and the distinction from the broad focus condition approaches significance in one of the two conditions with multiple focus.

Apart from the tendency just mentioned, these results on multiple foci hold equally of our two multiple focus conditions, in which the multiple focus is triggered by a multiple question (condition F-SU-M-q) and by multiple contrast (condition F-SU-M-c).

These findings support the prosodic account in (15), which is repeated here as (20). In this account it is not the feature [F] that triggers phonetic consequences (F0 raising and later lowering/compression). Rather, according to the indirect reference hypothesis, the syntactic feature [F] affects only the prosodic structure (not other phonology, nor phonetics). The effect we assume is stated in (12). The consequence is that for the complex focus only a single peak of intonation phrase stress is assigned. Given the choice among the two foci where this peak is assigned, we assume that (5) makes a decision in favor of the rightmost focus. It is then this abstract beat of intonation phrase stress that seems to be responsible for triggering the phonetic effects of F0-raising in the prominent syllable and F0-lowering/compression in later syllables.

(20) Expectation of the prosodic account (12): one i-phrase prominence for complex focus: phonetic raising and following lowering/compression for the i-phrase prominence

\[
\text{[subject]}_{F} \text{ verb } \left(\text{[modifier]}_{F} \text{ noun}\right)_{\text{object}}
\]

intonation phrase level prominence

We think that our results are not compatible with the view that each [F]-feature directly triggers F0-raising and subsequent lowering/compression, as shown in (21), repeated from (16).

(21) Expectation of a direct phonetic account: phonetic raising and following lowering/compression for each [F]-marked constituent

\[
\text{[subject]}_{F} \text{ verb } \left(\text{[modifier]}_{F} \text{ noun}\right)_{\text{object}}
\]

Our results also suggest that the phonetic focus effect of Mandarin Chinese would not be properly modeled by a prosodic account that requires intonation phrase prominence for each focus separately in multiple-focus constructions of the kind investigated here. Such an account would also lead us to expect the F0-pattern in (16). In this regard, the distinction drawn between (10) and (11) above seems to be appropriate. The structures investigated here seem to be parallel to (10), with a single scope for the multiple foci, and the assignment of a single location of intonation phrase level prominence in multiple focus constructions by (12).

There is no effect of global F0 raising due to greater speaker involvement (Bolinger, 1986) in the focus conditions apparent in our data: This would have raised the F0 height in the strings preceding a narrow focus on the modifier; however, there is no
such raising. Furthermore, if a lower amount of prominence than intonation phrase prominence is assigned to the first of two foci by constraints like (12), then, on the prosodic account, this lower amount of prominence would seem to be below the threshold that has a clearly noticeable effect on $F_0$. At the same time, the greater amount of prominence on the second focus that is expected by (12) is strong enough to have the described effect on $F_0$ height. The prosodic account in (20) is compatible with the possibility that one or another of these factors are present in the Dutch data of Rump and Collier (1996) and in the English data of Eady et al. (1986). The Mandarin Chinese data would then have allowed us to isolate and separately inspect the effect of the intonation phrase level prominence peak.

We think that our data strongly suggests that it is not the case that each [F]-feature directly triggers raising in [F] and following lowering and compression. However, a proponent of the direct phonetic account may formulate such a revised account that is compatible with our data. Assume that the two foci are construed as two parts of an abstract entity, a complex focus consisting of these two F-marked elements. One might now maintain that $F_0$-lowering/compression is triggered in a position that follows this complex focus, and thus occurs after the last one of them. One might further maintain that the effect of raising in the focus is distributed across the two elements of the complex focus. Instead of a larger amount of raising in a single focus, there may be smaller amounts of raising in each of the two foci. So long as these are small enough, they could be below the significance threshold in our data, but not falsified by it. To be sure, our data also do not directly provide evidence for such an alternative. A more extensive study would be required to distinguish such more subtle predictions.

4.2. Raven vs. winter melon

Why does sentence 1 show the focus peak on the second syllable of the modifier, while sentence 2 shows the focus peak on the first syllable of the modifier? In both cases, lowering after the focus sets in with the syllable following this peak. The syllables on which the focus peaks occur are underlined in (22).

(22)  a. [Whose nest does a kitty steal?]
    māomī tōu [wūyā]₄ wō.
    (Sentence 1)
    ‘A kitty steals a nest’

   b. [What soup does the aunt cook?]
    Gūmā shǎo [dōngguā]₄ tāng.
    (Sentence 2)
    ‘The aunt cooks a winter melon soup.’

A possibility suggested by a reviewer is that the second syllable of dōngguā, ‘winter melon’ carries a neutral tone. However, this is not compatible with our other measurements. Neutral tone is clearly discernible in our measurements: It can be seen in Fig. 3 that the focus-neutral condition (thick black line) of sentences 3 and 4 shows a dramatic $F_0$-dip due to the neutral tone in position N6. The utterance it occurs in is otherwise composed of only H-toned syllables. This focus-neutral condition would, without the neutral tone, follow a more or less horizontal course, as the corresponding condition in sentence 1 of Fig. 3. Given that neutral tone shows such a clear $F_0$-effect in the environment of H tones, we can assess whether dōngguā, ‘winter melon’ contains a neutral tone in the focus-neutral condition of sentence 2 of Fig. 3. Sentence 2, which contains dōngguā, ‘winter melon’ in the positions plotted as H4 and H5, should show a dramatic dip in the position H5 of the focus-neutral condition (thick black line), if it contained a neutral tone there. This is clearly not the case. Thus, it seems that dōngguā, ‘winter melon’, does not contain a neutral tone. The early peak on H4 rather than H5 of sentence 2, when this modifier is in focus, has a different reason.

We do not have a conclusive analysis, but the following assessment seems to us to be reasonable. We adopt from Duanmu (2007) that compounds and multi-syllabic words have a syllable that carries the strongest prominence in them in Mandarin Chinese. Whether this is the first or second syllable in a two-syllable word depends on various factors according to Duanmu. These include the morphological headedness of a compound, its frequency (with treatment as a non-compound word in cases of frequent collocations) and whether another word is following. Duanmu’s suggestions do not allow us to predict that dōngguā, ‘winter melon’ is different from the other two-syllable words we employed. However, it seems reasonable to maintain that the difference we observe is one in terms of prominence-assignment in two-syllable words or compounds. As shown in (23), this prosodic analysis in terms of prominence fits nicely with the analysis of the focus effects in terms of prominence: The syllable that carries the strongest stress in the two-syllable words or compounds is the one that is strengthened by the grid-mark assigned at the level of the intonation phrase due to focus.

(23)  a. x intonation phrase level prominence due to F

   x word- or phrase-level prominence
    gūmā shǎo [dōngguā]₄ tāng

   b. x intonation phrase level prominence due to F

   x word- or phrase-level prominence
    māomī tōu [wūyā]₄ wō

E. Kabagema-Bilan et al. / Lingua 121 (2011) 1890–1905
The assumption that the assignment of higher-level prominence normally amounts to strengthening a local prominence-maximum at a lower level is standard in accounts of prominence, with abundant empirical justification. Liberman and Prince (1977:250f) call this the “preservation of relative prominence under embedding”. For stress-representations in terms of metrical grids as in (23), this is sometimes cast in terms of the Continuous Column Constraint (see, e.g. Hayes, 1995:34ff): The higher grid-mark in (23a) could not be assigned to the second syllable of dōnggua, since it would not stand on a lower grid-mark there.

In sum, we think it is reasonable to hold a prosodic difference in word or compound stress responsible for the different peak locations for the different words in sentences 1 and 2 in our experiment. Notice that, if this analysis of different focus peak locations in Mandarin Chinese can be shown to be correct for a greater variety of words, it provides strong support for the prosodic analysis of the phonetic focus effects in this language. The prosodic analysis of the focus effects correctly predicts that the location of the intonation phrase peak is sensitive to prominence positions at lower prosodic levels. If, on the other hand, F0 raising and later lowering and compression is a direct phonetic effect of the feature [F], it would apply to the F-marked constituent qua F-marking, and the timing of the peak would not be linked to the position of greatest prominence within that constituent.\

5. Conclusion

Raising in the focus and subsequent lowering and compression of the F0 range are the phonetic F0-correlates of a single focus in Mandarin Chinese. In our experiment on multiple focus, we have found that a sequence of two foci does not show this effect on both foci, but only on the second focus (where lowering following the second focus was clearer than raising in the second focus).

We have shown that this is compatible with an account in which the relation between focus and its phonetic correlates is mediated by abstract intonation phrase prominence: Each focus may require some prominence (Stress-F in (12b)), but in cases of multiple focus it is sufficient if one intonation phrase prominence is assigned (Focus in (12a)). We assume that this is assigned to the rightmost focus, in line with a cross-linguistic tendency for rightmost prominence at the level of the intonation phrase. Our results do not seem to be compatible with an account in which each F-feature triggers raising and subsequent lowering.

We also encountered a difference in focus intonation between narrow focus on wūyā, ‘raven’, with a focus peak on the second syllable and narrow focus on dōnggua, ‘winter melon’, with a focus peak on the first syllable. If we are correct in relating this to prominence at lower prosodic levels, this distinction further supports our account in which the phonetic consequences of focus are mediated by a representation of abstract prominence.

Acknowledgements

This work was supported by the German Science Foundation (DFG) in connection with the SPP 1234, ‘Phonological and phonetic competence: between grammar, signal processing, and neural activity’, project TR747-2, and by the German Federal Ministry of Education and Research (BMBF), Grant Nr. 01UG0711. The names of the authors are arranged in alphabetical order.

Appendix A

The four target sentences in five focus conditions.

<table>
<thead>
<tr>
<th>Questions</th>
<th>Target sentences</th>
</tr>
</thead>
<tbody>
<tr>
<td>有什么新消息？</td>
<td>1. 猫咪偷乌鸦窝。</td>
</tr>
<tr>
<td>Yǒu shénme xīn xiàoxí?</td>
<td>Māomì tōu wūyā wō.</td>
</tr>
<tr>
<td>‘What’s new?’</td>
<td>‘A kitty steals a raven nest.’</td>
</tr>
<tr>
<td>谁偷谁的窝？</td>
<td>2. 谁偷谁的鸟。</td>
</tr>
<tr>
<td>Shuí tōu shuí de wō?</td>
<td>谁偷谁的鸟。</td>
</tr>
<tr>
<td>‘Who steals whose nest?’</td>
<td>‘Who steals whose nest?’</td>
</tr>
<tr>
<td>老鹰偷喜鹊窝吗？</td>
<td>3. 老鹰偷喜鹊窝吗？</td>
</tr>
<tr>
<td>Lǎoyīng tōu xǐquè wō ma?</td>
<td>‘Does an eagle steal a magpie’s nest?’</td>
</tr>
</tbody>
</table>

3 A reviewer brings up the possibility that [F] serves as the trigger for a process of raising F0 in the position of greatest word-level prominence. In terms of the model in (1), this would be an instance in which a syntactic feature ([F]) and a prosodic (phonological) element (word-level prominence) jointly define a phonetic effect (F0-raising in the location of the latter). In our view, this would be highly unexpected in a modular conception like that in (1).
谁偷乌鸦窝？
Shéi tōu wūyā wō?
'Who steals a raven's nest?'

猫咪偷谁的窝？
Māomì tōu shéi de wō?
'Whose nest steals does a kitty steal?'

有什么新消息？
Yōu shénme xīn xiāoxi?
'What's new?'

谁烧什么汤？
Shéi shāo shénme tāng?
'Who cooks what soup?'

伯父烧番茄汤吗？
Bófu shāo fánqié tāng ma?
'Does an uncle cook a tomato soup?'

谁烧冬瓜汤？
Shéi shāo dōngguā tāng?
'Who cooks a winter melon soup?'

姑妈烧什么汤？
Gūmā shāo shénme tāng?
'What soup does the aunt cook?'

有什么新消息？
Yōu shénme xīn xiāoxi?
'What's new?'

谁拎谁的书包？
Shéi līn shéi de shūbāo?
'Who carries whose school bag?'

李四拎张三的书包吗？
Lǐ Sì līn ZhāngSān de shūbāo ma?
'Does Li Si carry ZhangSan’s school bag?'

谁拎汪优的书包？
Shéi līn Wāng Yōu de shūbāo?
'Who carries Wang You’s school bag?'

张巍拎谁的书包？
Zhāng Wēi līn shéi de shūbāo?
'Whose school bag does Zhang Wei carry?'

有什么新消息？
Yōu shénme xīn xiāoxi?
'What's new?'

谁摸谁的猫咪？
Shéi mō shéi de māomī?
'Who strokes whose kitty?'

张三摸李四的猫咪吗？
Zhāng Sān mō LǐSì de māomī ma?
'Does Zhang San stroke Li Si’s kitty?'

谁摸潘安的猫咪？
Shéi mō Pān Ān de māomī?
'Who strokes Pan An’s kitty?'

孙荫摸谁的猫咪？
Sūn Yīn mō shéi de māomī?
'Whose kitty does Sun Yin stroke?'

2. 姑妈烧冬瓜汤。
Gūmā shāo dōngguā tāng.
'The aunt cooks a winter melon soup.'

3. 张巍拎汪优的书包。
Zhāng Wēi līn Wāng Yōu de shūbāo.
'Zhang Wei carries Wang You's school bag.'

4. 孙荫摸潘安的猫咪。
Sūn Yīn mō Pān Ān de māomī.
'Sun Yin strokes Pan An’s kitty.'
References


