



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

ARTICLE INFO

Article history:

Received 3 May 2009

Received in revised form 19 January 2011

Accepted 11 February 2011

Available online 1 April 2011

Keywords:

Focus

Multiple focus

Intonation

Prosody

Mandarin Chinese

ABSTRACT

Xu (1999) reported that a single (non-sentence-final) focus in Mandarin Chinese has the phonetic effect of F_0 raising in the focus and F_0 lowering and compression following the focus. In this paper, results of an experiment on multiple (non-sentence-final) focus in Mandarin Chinese are reported, with baseline conditions single focus and no focus. The single focus conditions showed the expected effects. In the multiple focus conditions, only the second of two foci shows phonetic F_0 effects, while the first focus does not show such F_0 effects. Duration measurements also provide no clear evidence for a phonetic focus effect on the first of the two foci. The analysis we suggest, combining suggestions of Schwarzschild (1999) and Selkirk (2005), is that one focus in a focus-construction attracts intonation phrase stress, and that the phonetic effects are triggered by intonation phrase stress, rather than by the F-feature that marks focus.

© 2011 Elsevier B.V. All rights reserved.

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 F_0 in the focus and lowers and compresses F_0 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

* Corresponding author. Tel.: +49 30 20192 424; fax: +49 30 20192 402.

E-mail addresses: lenabilan@yahoo.com (E. Kabagema-Bilan), beaberbe@hotmail.com (B. López-Jiménez), truckenbrodt@zas.gwz-berlin.de (H. Truckenbrodt).

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 | intonation phrase level |
| x | x | phonological phrase level |
- ... 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)
- | | | | | | | | |
|---|--|---|-----|---|--|---|-----|
| (| | x |)IP | (| | x |)IP |
| I suggest that the crime was committed [in the lounge] _F [by Mr. Green] _F | | | | | | | |
| (| | | | | | | |
| x | | | | | | | |
|)IP | | | | | | | |
| [with a wrench] _F . | | | | | | | |

It seems to us that there is an empirically real difference between examples of the kind used by Schwarzschild and examples of the kind considered by Selkirk. In German, examples of the kind in (6), in their context, are likewise naturally produced with larger divisions and intonation phrase stress, as shown in (7).

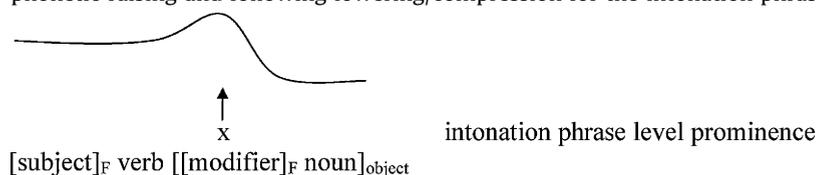
- (7)
- | | | | |
|---|--|---|-----|
| (| | x |)IP |
| Ich schlage vor, dass das Verbrechen im Salon begangen wurde, | | | |
| I suggest that the crime in the lounge committed was, | | | |
| (| | x |)IP |
| von Mr. Green, (und) mit einer Rohrzanze. | | | |
| by Mr. Green and with a wrench | | | |

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) [Who read what to whom?]
- | | | | | |
|---|---|---|----------------------------------|--------------------------------|
| (| | x |)IP | intonation phrase level stress |
| x | x | x | phonological phrase level stress | |
| [Hans] _F hat [Maria] _F [die Zeitung] _F vorgelesen, | | | | |
| Hans has Maria the newspaper read | | | | |
| (| | x |)IP | intonation phrase level stress |
| x | x | x | phonological phrase level stress | |
| und [Peter] _F hat [Claudia] _F [ein Buch] _F vorgelesen. | | | | |
| and Peter has Claudia a book read | | | | |

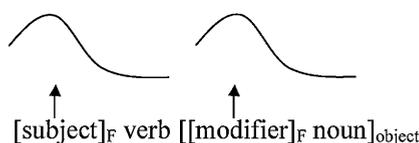
‘Hans read the newspaper to Maria, and Peter read a book to Claudia.’

- (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



A direct account of the phonetic consequences of [F] makes markedly different predictions, as shown in (16). If the feature [F] directly triggers phonetic F_0 -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



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 F_0 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 F_0 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 F_0 -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 F_0 -height. (c) Bolinger (1986) suggested that a greater amount of speaker involvement raises F_0 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 F_0 -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 F_0 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 F_0 raising and later lowering/compression.

We now turn to our experiment on multiple focus in Mandarin Chinese.

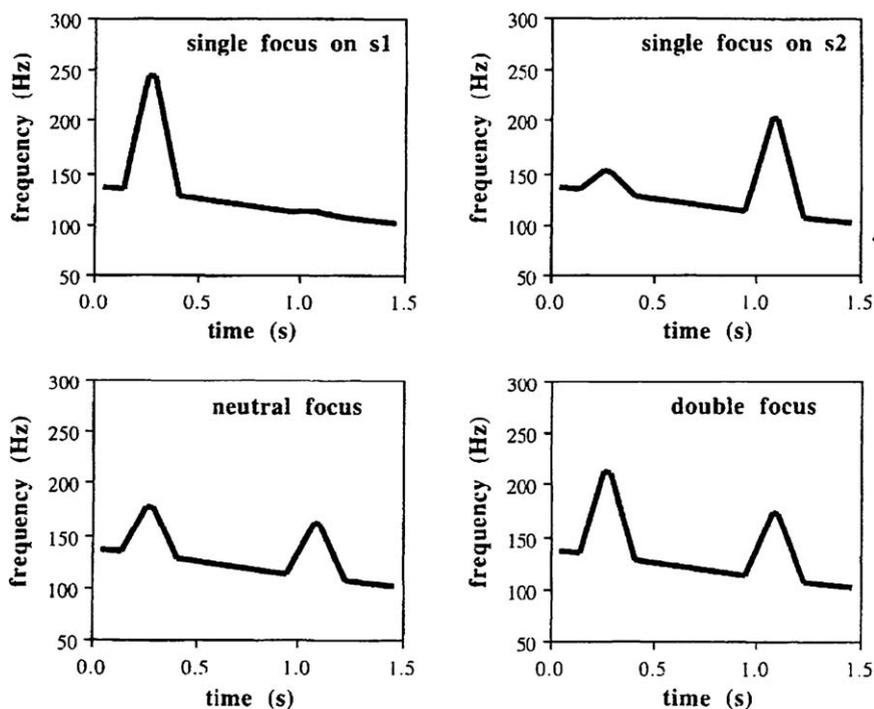


Fig. 2. Averaged adjustments (prototypical contours) reported by Rump and Collier (1996:9).

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 F_0 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

	Subject	Verb	Modifier	Gram. particle	Object noun
Sentence 1	māomī 'kitty'	tōu 'steal'	wūyā 'raven'		wō 'nest'
Sentence 2	gūmā 'aunt'	shāo 'cook'	dōngguā 'winter melon'		tāng 'soup'
Sentence 3	Zhāng Wēi 'Zhang Wei' (name)	līn 'carry'	WāngYōu 'Wang You' (name)	de 's	shū-bāo 'school bag'
Sentence 4	Sūn Yīn 'Sun Yin' (name)	mō 'stroke'	PānĀn 'Pan An' (name)	de 's	māomī 'kitty'
	Focus 1		Focus 2		

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).¹

(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

1. Broad	Yǒu shénme xīn xiāoxi?	‘What’s new?’
2. F-SU-M-q	Shéi tōu shéi de wō?	‘Who steals whose nest?’
3. F-SU-M-c	Lǎoyīng tōu xǐquè wō ma?	‘Does an eagle steal a magpie’s nest?’
4. F-SU	Shéi tōu wūyā wō?	‘Who steals a raven’s nest?’
5. F-M	Māomī tōu shéi de wō?	‘Whose nest does a kitty steal?’
Target sentence for 1.–5.	māomī 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 province² 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 *ZhāngWēi* (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 F_0 values were measured at these points. In a number of utterances, the F_0 value of the sentence final syllables (the last two to three syllables) in postfocal position could not be measured due to creaky voice.

The F_0 -measurements were normalized for speaker-specific tonal height. For each speaker *S*, the average height of the first syllable in unfocused utterances, call it $H1_s$, is first computed. The normalized values of this speaker are obtained by dividing each values of this speaker through the value of $H1_s$. 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. F_0 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.

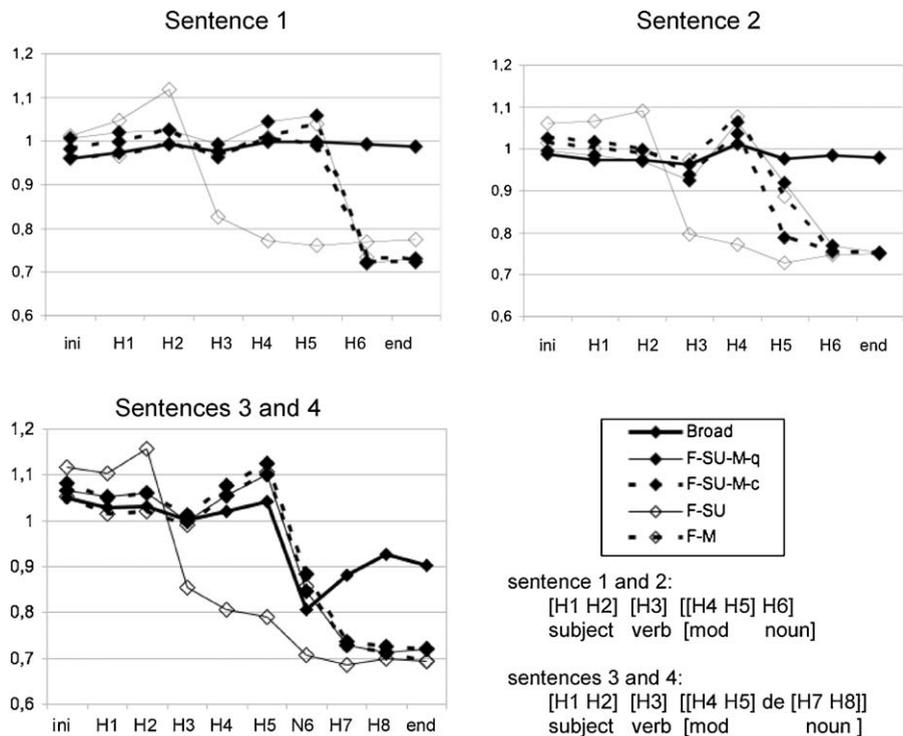


Fig. 3. Pooled normalized results for the five focus conditions of the four target sentences. 1. *Māomī tōu wūyǎ wō.* ‘A kitty steals a raven nest.’ 2. *Gūmā shāo dōngguā tāng.* ‘The aunt cooks a winter melon soup.’ 3. *Zhāng Wēi lǐn Wáng Yōu de shūbāo.* ‘Zhang Wei carries Wang You’s school bag.’ 4. *Sūn Yīn mō Pān Ān de māomī.* ‘Sun Yin strokes Pan An’s kitty.’

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 F_0 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 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.

	H2 (subject)	H3 (verb)	H5/H4 (object modifier)	H6/H8 (last syll.)
df, <i>F</i>	df = 4; <i>F</i> = 14,1	df = 4; <i>F</i> = 36,9	df = 4; <i>F</i> = 89,9	df = 4; <i>F</i> = 53,8
Sig	$p < 0.001$	$p < 0.001$	$p < 0.001$	$p < 0.001$
Post hoc bearing on single <i>F</i>	F-SU > Broad F-SU > F-M	Broad > F-SU F-M > F-SU	F-M > Broad Broad > F-SU F-M > F-SU F-SU-M-q > F-SU F-SU-M-c > F-SU	Broad > F-SU Broad > F-M
Post hoc bearing on two <i>F</i>	F-SU > F-SU-M-q F-SU > F-SU-M-c	F-SU-M-q > F-SU F-SU-M-c > F-SU	(F-SU-M-q ≥ Broad; $p = .054$)	Broad > F-SU-M-q Broad > F-SU-M-c
Other sign. dist.	None	None	None	None

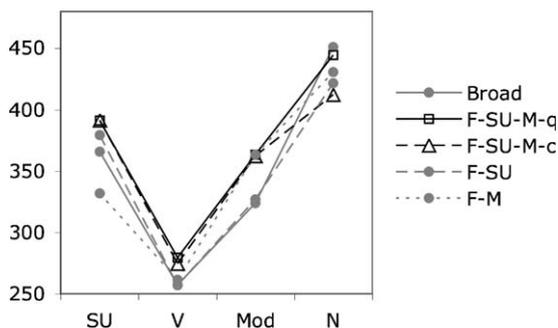


Fig. 4. Duration measurements.

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 F_0 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 2
Significance results of duration measurements.

	SU	V	Mod
Single <i>F</i> vs. Broad			F-M > Broad F-M > F-SU
Multiple focus	F-SU-M-c > Broad F-SU-M-c > F-M F-SU-M-q > F-M	F-SU-M-c > Broad F-SU-M-q > Broad F-SU-M-c > F-SU F-SU-M-q > F-SU F-SU-M-q > F-M	F-SU-M-c > Broad F-SU-M-q > Broad F-SU-M-c > F-SU F-SU-M-q > F-SU
Other sig. dist.	All other conds. > F-M	None	None

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 F_0 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.

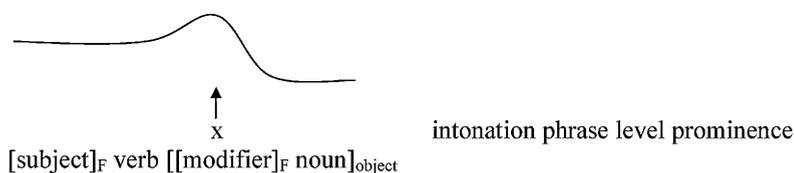
The first of two foci on the sentence subject is statistically not distinct from the subject in the broad focus condition. The following verb (where lowering might occur) is also non-distinct in F_0 in the two conditions. At the same time, the sentence subject has significantly higher F_0 when it alone 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 F_0 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).

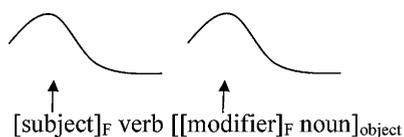
These findings support the prosodic account in (15), which is repeated here as (20). In this account it is not the feature [F] that triggers phonetic consequences (F_0 raising and later lowering/compression). Rather, according to the indirect reference hypothesis, the syntactic feature [F] affects only 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 F_0 -raising in the prominent syllable and F_0 -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



We think that our results are not compatible with the view that each [F]-feature directly triggers F_0 -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



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 constructions of the kind investigated here. Such an account would also lead us to expect the F_0 -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 F_0 raising due to greater speaker involvement (Bolinger, 1986) in the focus conditions apparent in our data: This would have raised the F_0 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 then 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]_F wō. (Sentence 1)
 'A kitty steals a raven's nest'
- b. [What soup does the aunt cook?]
 Gūmā shāo [dōngguā]_F 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 shāo [dōngguā]_F tāng
- b. x intonation phrase level prominence due to F
 x word- or phrase-level prominence
 māomī tōu [wū yā]_F wō

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 *dōngguā*, 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, F_0 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.³

5. Conclusion

Raising in the focus and subsequent lowering and compression of the F_0 range are the phonetic F_0 -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 *dōngguā*, ‘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.

Questions	Target sentences
有什么新消息? Yǒu shénme xīn xiāoxi? ‘What’s new?’	1. 猫咪偷乌鸦窝。 Māomī tōu wūyā wō. ‘A kitty steals a raven nest.’
谁偷谁的窝? Shéi tōu shéi de wō? ‘Who steals whose nest?’	
老鹰偷喜鹊窝吗? Lǎoyīng tōu xǐquè wō ma? ‘Does an eagle steal a magpie’s nest?’	

³ A reviewer brings up the possibility that [F] serves as the trigger for a process of raising F_0 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 (F_0 -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？

Māomī tōu shéi de wō?

'Whose nest steals does a kitty steal?'

有什么新消息？

Yǒu 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？

Bófù shāo fānqié 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？

Yǒu 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？

Yǒu 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

- Baumann, S., Grice, M., Steindamm, S., 2006. Prosodic marking of focus domains—categorical or gradient? *Proceedings of Speech Prosody* 301–304.
- Baumann, S., Becker, J., Grice, M., Mücke, D., 2007. Tonal and articulatory marking of focus in German. *ICPhS* 1029–1032.
- Bolinger, D., 1986. *Intonation and Its Parts: Melody in Spoken English*. Stanford University Press, Palo Alto.
- Büring, D., 2009. Towards a typology of focus realization. In: Zimmermann, M., Féry, C. (Eds.), *Information Structure*. Oxford University Press, Oxford, pp. 177–205.
- Chen, Y., Gussenhoven, C., 2008. Emphasis and tonal implementation in Standard Chinese. *Journal of Phonetics* 36, 724–746.
- Duanmu, S., 2007. *The Phonology of Standard Chinese*, second edition. Oxford University Press, Oxford.
- Eady, S.J., Cooper, W.E., Klouda, G.V., Mueller, P.R., Lotts, D.W., 1986. Acoustical characteristics of sentential focus: narrow vs. broad and single vs. dual focus environments. *Language and Speech* 29, 233–250.
- Gårding, E., 1987. Speech Act and Tonal Pattern in Standard Chinese: Constancy and Variation. *Phonetica* 44, 13–29.
- Gussenhoven, C., 1992. Sentence accents and argument structure. In: Roca, I. (Ed.), *Thematic Structure, Its Role in Grammar*. Foris, Berlin, New York, pp. 79–106.
- Gussenhoven, C., 2004. *The Phonology of Tone and Intonation*. Cambridge University Press, Cambridge.
- Hayes, B., 1995. *Metrical Stress Theory: Principles and Case Studies*. The University of Chicago Press, Chicago.
- Hirschberg, J., Ward, G., 1992. The influence of pitch range, duration, amplitude and spectral features on the interpretation of the rise-fall-rise intonation contour in English. *Journal of Phonetics* 20, 241–251.
- Inkelas, S., 1989. *Prosodic constituency in the lexicon*. Doctoral Dissertation. Stanford University.
- Jackendoff, R.S., 1972. *Semantic Interpretation in Generative Grammar*. MIT Press, Cambridge, MA.
- Jin, S., 1996. *An acoustic study of sentence stress in Mandarin Chinese*. Doctoral Dissertation. Ohio State University.
- Katz, J., Selkirk, E., 2009. Contrastive Focus vs. Discourse-new: Evidence from Prosodic Prominence in English. Ms., MIT and UMass, Amherst.
- Kochanski, G., Shih, C., 2003. Prosody modeling with soft templates. *Speech Communication* 41, 625–645.
- Ladd, D.R., 2008. *Intonational Phonology*, second edition. Cambridge University Press, Cambridge.
- Liberman, M., Prince, A., 1977. On stress and linguistic rhythm. *Linguistic Inquiry* 8, 249–336.
- Liu, F., Xu, Y., 2005. Parallel encoding of focus and interrogative meaning in Mandarin intonation. *Journal of Phonetics* 33, 70–105.
- Pierrehumbert, J.B., 1980. *The phonology and phonetics of English intonation*. Doctoral Dissertation. MIT.
- Rooth, M., 1992. A theory of focus interpretation. *Natural Language Semantics* 1, 75–116.
- Rooth, M., 1996. Focus. In: Lappin, S. (Ed.), *The Handbook of Contemporary Semantic Theory*. Blackwell, Oxford and Cambridge, MA, pp. 271–297.
- Rump, H.H., Collier, R., 1996. Focus conditions and the prominence of pitch-accented syllables. *Language and Speech* 39, 1–17.
- Schwarzschild, R., 1999. Givenness, AvoidF and other constraints on the placement of accent. *Natural Language Semantics* 7, 141–177.
- Selkirk, E., 1995. Sentence prosody: intonation, stress, and phrasing. In: Goldsmith, J. (Ed.), *The Handbook of Phonological Theory*. Blackwell, Cambridge, MA, pp. 550–569.
- Selkirk, E., 2002. Contrastive FOCUS vs. presentational focus: prosodic evidence from right node raising in English. *Proceedings of Speech Prosody* 643–646.
- Selkirk, E., 2004. Bengali intonation revisited: an optimality theoretic analysis in which FOCUS stress prominence drives FOCUS phrasing. In: Lee, C.-M. L., Gordon, M., Buerling, D. (Eds.), *Topic and Focus: A Cross-Linguistic Perspective*. Kluwer, Dordrecht, pp. 217–246.
- Selkirk, E., 2005. Comments on intonational phrasing in English. In: Frota, S., Vigário, M., Freitas, M.J. (Eds.), *Prosodies. With Special Reference to Iberian Languages*. Mouton de Gruyter, Berlin, New York, pp. 11–58.
- Shih, C., 1988. Tone and Intonation in Mandarin. In: *Working Papers of the Cornell Phonetics Laboratory No. 3. Stress, Tone and Intonation*. Department of Linguistics, Cornell University, Ithaca, NY, pp. 83–108.
- Shih, C., 2004. Tonal effects on intonation. In: *International Symposium on Tonal Aspects of Languages: With Emphasis on Tone Languages*, ISCA archives, pp. 163–168.
- Sugahara, M., 2005. Post-focus prosodic phrase boundaries in Tokyo Japanese: asymmetric behavior of an F0 cue and domain-final lengthening. *Studia Linguistica* 59, 144–173.
- Truckenbrodt, H., 1995. *Phonological phrases: their relation to syntax, focus, and prominence*. Doctoral Dissertation. MIT.
- Truckenbrodt, H., 2006. Phrasal stress. In: Brown, K. (Ed.), *The Encyclopedia of Languages and Linguistics*, second edition, vol. 9. Elsevier, Oxford, pp. 572–579.
- Truckenbrodt, H., 2007. The syntax–phonology interface. In: de Lacy, P. (Ed.), *The Cambridge Handbook of Phonology*. Cambridge University Press, Cambridge, pp. 435–456.
- Truckenbrodt, H. Checking vs. stressing of wh-words. In: Elordieta, G., Barbosa, P. (Eds.), *Prosody and Meaning*. Mouton, Berlin, in press.
- Xu, Y., 1999. Effects of tone and focus on the formation and alignment of f0 contours. *Journal of Phonetics* 27, 55–105.
- Yuan, J., 2004. *Intonation in Mandarin Chinese: acoustics, perception and computational modelling*. Doctoral Dissertation. Cornell University.