Original Paper

Phonetica 2018;75:24–56
DOI: 10.1159/000453082

Received: November 13, 2016
Accepted after revision: November 5, 2016
Published online: June 9, 2017

Interactive Prosodic Marking of Focus, Boundary and Newness in Mandarin

Bei Wang[a] Yi Xu[b] Qifan Ding[c]

[a]Institute of Chinese Minority Language and Literatures, Minzu University of China, Beijing, China; [b]Department of Speech, Hearing and Phonetic Sciences, University College London, London, UK; [c]Princeton University, East Asian Department, Princeton, NJ, USA

Abstract

The current study investigates whether and how focus, phrase boundary and newness can be simultaneously marked in speech prosody in Mandarin Chinese. Homophones were used to construct three syntactic structures that differed only in boundary condition, focus was elicited by preceding questions, while newness of postboundary words was manipulated as whether they had occurred in the previous text. Systematic analysis of F_0 and duration showed that (1) duration was a reliable correlate of boundary strength regardless of focus location, while involvement of F_0 was only in terms of lowering of phrase-final F_0 minima and raising of phrase-initial F_0 minima at a relatively strong boundary, (2) postfocus compression (PFC) of F_0 was applied across all boundaries, including those with long silent pauses (over 200 ms), and postfocus F_0 was lowered to almost the same degree in all boundary conditions, and (3) newness of postfocus words had no systematic effect on F_0 or duration. These results indicate that not only functionally focus is independent of prosodic structure and newness, but also phonetically its realization is separate from boundary marking. Focus is signaled mainly through pitch range adjustments, which can occur even across phrase breaks, whereas boundaries are mostly signaled by duration adjustments.

1 Introduction

Focus is a semantic and pragmatic function that can be realized via prosody to highlight a particular constituent in an utterance (Bolinger, 1958; Eady and Cooper, 1986; Xu, 1999, 2005). There has been rich empirical evidence that in many languages focus is realized mainly by increasing the pitch range, intensity, duration, and articulatory fullness of the focused word, and reducing the F_0 and intensity of the following words, while leaving the prefocus words largely unchanged (English: Cooper et al., 1985; de Jong, 1995; Xu and Xu, 2005; Mandarin: Chen and Gussenhoven,

KARGER

E-Mail karger@karger.com
www.karger.com/pho

Yi Xu
Department of Speech, Hearing and Phonetic Sciences
University College London, Chandler House
2 Wakefield Street, London WC1N IPF (UK)
E-Mail yi.xu@ucl.ac.uk

2008; Wang and Xu, 2011; Xu, 1999; German: Féry and Kügler, 2008; Greek: Botinis et al., 1999; Dutch: Swerts et al., 2002; Japanese: Ishihara, 2002; Korean: Lee and Xu, 2010; Turkish: Ipek, 2011; Tibetan: Wang et al., 2012; Zhang et al., 2012; Estonian: Sahkai et al., 2013). The reduction of postfocus F_0 and intensity is known as postfocus compression or PFC (Xu et al., 2012), and it is found to be critical for focus perception in at least some of these languages (Vainio et al., 2003, for Finnish; Rump and Collier, 1996, for Dutch; Prom-on et al., 2009, for English; Ishihara, 2011, and Sugahara, 2005, for Japanese; Liu and Xu, 2005, and Xu et al., 2012, for Mandarin). What is yet unclear, and in fact rarely asked, is how extensive the temporal domain of PFC is. The empirical works on focus just mentioned are all done in relatively simple sentences. Those sentences are said with a single breath without noticeable or consistent pauses. Thus, it is still unknown whether PFC can be blocked by a prosodic boundary with an apparent pause. This issue is, however, closely related to many other issues that have been the center of both theoretical and empirical examinations, including, in particular, boundary marking and newness. In the following we will first briefly review the literature on these issues as they are related to focus before outlining our specific research questions.

1.1 Boundary Marking

Boundary marking has been a controversial topic, with disagreement both between theories in terms of boundary types and levels, and between empirical findings in terms of acoustic cues of boundary.

A widespread notion is that there exists a prosodic structure in speech in the form of a hierarchy (Beckman, 1996; Selkirk, 1986). This structure exists in its own right and is largely autonomous from syntactic structures as there are often mismatches between the two (Ladd, 2008). In this structure the largest unit is what is known as intonational phrase (Beckman and Pierrehumbert, 1986; Ladd, 2008; Pierrehumbert, 1980), tone group (Halliday, 1967; O'Connor and Arnold, 1961; Palmer, 1922; Wells, 2006), or intonation group (Cruttenden, 1997). Critically, the temporal domain of this top unit is defined in terms of its internal constituent. In the nuclear tone analysis tradition, a tone group is said to consist of an obligatory nuclear tone and an optional head, prehead and tail (O'Connor and Arnold, 1961; Palmer, 1922; Wells, 2006). In the autosegmental-metrical theory of intonation, an intonational phrase is defined as consisting of an obligatory nuclear accent and a boundary tone which is either high or low (Beckman and Pierrehumbert, 1986; Ladd, 2008; Pierrehumbert, 1980).

There is little agreement on the categories below the intonational phrase, however. Some have proposed that just one level in between is enough, e.g., phonological phrase (English: Nespor and Vogel, 1986; Turkish: Ipek and Jun, 2013; Japanese: Ishihara, 2011). Some are in favor of two levels, e.g., major phonological phrase and minor phonological phrase (Japanese: Kubozono, 1993, and Sugahara, 2005; English: Selkirk, 2005, and Selkirk et al., 2004), which are roughly equivalent to intermediate phrase and pitch accent phrase as proposed by Beckman and Pierrehumbert (1986). More recently, Ito and Mester (2013) proposed that major and minor phrase can be integrated into one category, namely φ-phrase, because they both serve as the domain of downstep and initial lowering. They further proposed that a φ-phrase can dominate another φ-phrase, forming a recursive structure. A minimal projection of a φ-phrase corresponds to a minor phrase, which contains at most one lexical pitch accent. In other words, the phrase domain is again defined in terms of its internal accentual or tonal components. Likewise, according to Selkirk (2005), "…the presence of a metrical prominence entails the presence of the constituent of which it is the head, there will be as many intonational phrase constituents as there are contrastive foci" (p. 17).

From the above review, it is clear that in theories of prosodic structure, domains and their boundaries are defined in terms of their inner distribution of tones or pitch accents. Thus, these prosodic-structure-based boundaries are not independent of pitch accent. That is, prosodic domains defined this way already have their internal pitch units known based on prior observations. Such theory-internal dependency therefore makes it virtually impossible to study the interaction of boundaries and pitch units.

There have been empirical works, however, that try to classify prosodic boundaries relatively independently of focus or tonal events. Two main methods have been used to provide either the initial division or the final classification of boundary types: by syntactic structure (Ladd, 1988; Lehiste, 1973; Lehiste et al., 1976) and by perceptual judgment of naïve listeners (de Pijper and Sandeman, 1994; Swerts, 1997; Wightman et al., 1992). The different boundaries are then compared in terms of various acoustic parameters. Although there is limited agreement on the exact prosodic categories across these studies, most of them found that boundary strength is related to F_0 (Beckman and Pierrehumbert, 1986; de Pijper and Sandeman, 1994; Nespor and Vogel, 1986; Swerts and Geluykens, 1994; Swerts et al., 1996), silent pause (de Pijper and Sandeman, 1994; Swerts, 1994, 1997, 1998), final lengthening (Ladd and Campbell, 1991; Oller, 1973; Swerts, 1994, 1997; Wightman et al., 1992), phrase-initial glottalization (Dilley et al., 1996), and phrase-initial strengthening (Cho and Keating, 2001; Cho et al., 2007; Fougeron, 2001; Fougeron and Keating, 1997; Jun, 1993).

However, there are disagreements on the role of F_0 in boundary marking. Some studies found boundary strength to be related to degree of preboundary F_0 lowering, postboundary F_0 raising and the sum of these two effects, i.e., pitch reset (Dutch: de Pijper and Sandeman, 1994, and Swerts, 1997; English: Ladd, 1988). Also it is widely held that boundary tones are used to mark the edges of utterances, i.e., a low boundary tone or low-ending F_0 as a cue for finality and a high boundary tone or high-ending F_0 as a cue for continuity (Swedish: Bruce, 1983; English and Japanese: Beckman and Pierrehumbert, 1986; English: Lieberman, 1967; Dutch: Swerts and Geluykens, 1994). But these meanings are also closely related to the statement/question contrast (Liu et al., 2013; Pierrehumbert and Hirschberg, 1990). Besides, F_0 is used to mark focus, tone, word stress, newness, and many other aspects of language. Thus, the direct role of F_0 in distinguishing boundaries can be clearly seen only when the other factors are effectively controlled.

In studies where all factors other than boundary strength are kept as constant as possible, F_0 is found to make little contribution to boundary marking, in contrast to the much more reliable contribution of duration (English: Allbritton et al., 1996; Lehiste, 1973; Lehiste et al., 1976; Korean: Jeon and Nolan, 2013; Mandarin: Xu and Wang, 2009). Lehiste (1973) and Lehiste et al. (1976) show that timing is the principal means by which structures such as "(The old men) (and women sat on the bench)" is distinguished from "(The old men and women) (sat on the bench)." Subsequent perception experiments have shown that duration alone is sufficient for syntactic disambiguation (English: Allbritton et al., 1996; Price, 1991) or perceptual determination of groupings in number strings (Korean: Jeon and Nolan, 2013). Xu and Wang (2009) compared various grouping patterns of words and phrases with 1–4 syllables in Mandarin

Chinese and found that syllable duration had the most consistent relation to grouping patterns, while corresponding variations in F_0 displacement could be explained by time pressure on tonal articulation resulting from duration shortening. Other studies that have found duration as reliable boundary cue include Duez (1982), Fant and Kruckenberg (1996), Krivokapic and Byrd (2012), Wagner (2005), and Wightman et al. (1992).

Thus, empirical studies have shown that it is important to examine the acoustic cues of boundary strength independently of other factors and to keep stimulus sentences as identical as possible while trying to manipulate boundary strength. In particular, syntactic structures can be used to control boundary strength directly. This makes it possible to examine the interaction of focus and boundary, because they are no longer conceptually confounded with each other. Note, however, that this method does not imply that there exists a one-to-one mapping between syntax and prosody, because it assumes only that some syntactic boundaries happen to have consistent prosodic marking, which is already empirically demonstrated, as discussed above (also see Selkirk, 2011, and Ito and Mester, 2013). Likewise, focus should be manipulated independently of other factors as well. Only through independent control can we reliably observe how focus and boundary interact with each other.

In fact, there have been both theoretical arguments and empirical evidence in support of separating boundary marking from pitch accent placement. Some have argued that speakers are free to place pitch accents on whichever part of an utterance they wish to highlight rather than being dictated by a prosodic structure (Bolinger, 1972; Chafe, 1974; Halliday, 1967). Recently, it has been shown that even for languages that have been argued to mark focus by changing phrase structure (Korean: Jun, 1993; Japanese: Nagahara, 1994, and Pierrehumbert and Beckman, 1988), no sign of phrasal marking for focus can be found once focus and phrasing are independently controlled (Ishihara, 2011; Kubozono, 2007; Lee and Xu, 2012). It has also been shown that durational adjustments for boundary and focus are largely independent of each other. Turk and Shattuck-Hufnagel (2007) have found that preboundary lengthening applies in both pitch-accented and unaccented phrase-final words in English. Likewise, Horne et al. (1995) have found that focus condition does not affect pauses in Swedish. For Mandarin Chinese, the lengthening of a syllable at an intonational phrase-initial position, as compared to an intonational medial position, occurs mostly in the onset, whereas the lengthening of a syllable in focus is relatively more global and spans over the whole focused constituent (Chen, 2004).

In summary, for the purpose of the current study, what we could learn from the above literature review is that boundary strength corresponds to syntactic structure, and that, although there is no one-to-one mapping, boundary strength could be marked independently from focus, and the role of F_0 on marking boundary is still controversial.

1.2 Focus and Newness

It is widely in consensus that focus is directly marked in intonation in many languages, including Mandarin Chinese (e.g., Xu, 1999). What remains controversial is whether newness is also marked directly via intonation and whether newness interacts with focus. The reason the present study takes newness into consideration is that when studying the domain of PFC, we need to figure out whether the newness of the postfocus region has any effect on its intonation, which may interact with PFC.

Newness has been said to be related to accentness, which indicates that newness is marked acoustically. Many have reported that *new* information is mostly accented in various languages while *given* information is mostly deaccented (Brown, 1983; Chafe, 1976; Féry and Kügler, 2008; Fowler and Housum, 1987; Halliday, 1967; Hirschberg, 1993; Ladd, 1996; Nooteboom and Kruyt, 1987; Nooteboom and Terken, 1982; Prince, 1981). But there have also been doubts that newness is clearly marked by intonation. For instance, Ladd (1996) pointed out that the alleged given/new contrast is confounded with accentuation, and when accentuation is removed, there is little acoustic difference between the two. Terken (1984) and Terken and Hirschberg (1994) noticed that apart from the given/new status, there are additional factors that also affect accent distribution. Wang and Xu (2011) found that, when focus and location of words in sentence were both systematically controlled in Mandarin, *given* and *new* had no F_0 difference under *in situ comparison* instead of *sequential comparison*. With in situ comparison, the newness of a word is manipulated by varying the preceding context while comparing the same word in the same sentence position. Sequential comparison, instead, checks the same word occurring at two different positions in one sentence in which the first occurrence is considered as new and the second as given. The disadvantage of sequential comparison is that word position is confounded with newness. Under in situ comparison, Wang and Xu (2011) found that the difference between given and new was only in duration, i.e., a new word was longer than a given word.

The more standard view these days is that given/new is independent of background/focus structure (e.g., Féry and Samek-Lodovici, 2006; Krifka, 2008; Selkirk, 2008). In other words, the choice of focus relies on pragmatic and semantic factors other than newness.

To keep the research question simple and straightforward, here we only discuss narrow focus, which is the most prominent part of a sentence for semantic and pragmatic reasons. The prosodic or acoustic realization of a focus is often referred to as accent, although under strict terminology, focus and accent are not the same. For the purpose of this paper, we will take direct acoustic measurements related to focus without interpreting them in terms of accents. We will do the same for newness by examining differences in F_0 and duration between new and given conditions without interpreting them in terms of accents. This allows us to treat focus and newness independently.

There is recent evidence that focus and newness involve different cognitive processes. Chen et al. (2014) recorded brain responses to focus and newness that have been independently controlled and found that focused words elicited a larger P2 and larger positivity than unfocused words. In contrast, new words elicited a larger N400 and a smaller late positive complex than given words. They concluded that the processing of focused words reflected attention allocation and immediate integration of focused information, whereas the processing of new words reflected difficulty in information integration or memory retrieval.

Thus, focus seems to serve a function of highlighting a particular constituent of an utterance to draw the listener's attention. It is therefore needed only from time to time when the speaker feels the need to highlight something in particular. As such it does not have to occur very often or even in every sentence. Newness, in contrast, is virtually ubiquitous, because speech is to convey information, which, by definition (Shannon, 1948), has to continuously offer contents that are newsworthy. As a result, although some cases of focus do coincide with newness, many others do not (Krifka, 2008). So far, however, we are not aware of a definition of focus that is precise enough to predict *all and only* actual occurrences of focus. Instead, there are only empirical paradigms that have been shown to reliably elicit focus, e.g. minidialogues that involve making corrections, or answering *wh* questions (Cooper et al., 1985; Féry and Kügler, 2008; Liu and Xu, 2005; Pell, 2001; Wang and Xu, 2011; Xu, 1999; Xu and Xu, 2005), etc.

Another related issue that we need to consider is focus type, because as described later in the material part, when we construct different contexts for controlling the newness of the postfocus region, it also involves different types of focus. It is widely believed that there are different types of focus, each with specific prosodic cues (Gussenhoven, 2007; Krifka, 2008; Selkirk, 2008). In particular, there is an assumed distinction between information (or presentational) and corrective (or contrastive) focus (Gussenhoven, 2007; Selkirk, 2002). The former occurs in the first sentence given below, while the latter occurs in the second. However, despite the plausibility of the semantic distinction, empirical findings so far have been mixed as to whether there are consistent acoustic differences between the two types of focus (Baumann et al., 2007; Hanssen et al., 2008; He et al., 2011; House and Sityaev, 2003; Hwang, 2012; Katz and Selkirk, 2011; Kügler and Ganzel, 2014; Sahkai et al., 2013):

Who said that? HELEN said that.
Did John say that? No, HELEN said that.

1.3 Aim of the Current Study

The literature review above has shown that it is possible to investigate the interaction between focus, newness, and boundary marking by controlling them independently. This means to avoid using prosodic structure to define pitch accent type (e.g., nuclear vs. prenuclear) and to avoid using pitch accent distribution to define the type and temporal scope of prosodic phrasing, or to use newness to define focus or vice versa. In the present study, we will examine the relationship between focus, newness, and boundary by eliciting focus with discourse contexts, manipulating newness with presence/absence of previous mentioning of the key words and controlling boundary strength with syntactic structures. Since focus, phrasing, and newness are all independently controlled, not only can their respective encoding mechanisms be studied, but also their interactions can be observed with minimal circularity. To make it clear, we take the following operational definitions of focus and newness.

"Focus" highlights a particular constituent in an utterance for pragmatic purposes, e.g., correction of the previous information or providing information asked by a *wh* question (Cooper et al., 1985; Wang and Xu, 2011; Xu, 1999).

"Newness" refers to a whole word/phrase that appears for the first time, whereas "givenness" refers to a whole or a part of word/phrase that has appeared in the prior context (Prince, 1992).

For boundary strength, we take the following operational definition. We will use three different syntactic structures to elicit three boundary strengths. Given that syntactic boundaries may not always have clear and prosodic markers, we here carefully chose sentences in which word boundary, phrase boundary and clause boundary are clearly distinguishable in prosody.

"Boundary strength" refers to how closely two constituents adhere to each other, which is determined by the size of syntactic constituents in this study. That is, word boundary, phrase boundary, and clause boundary correspond to weak, medium, and strong boundaries, respectively.

The specific research questions to be addressed are as follows:

(a) Does PFC apply across boundaries of different strengths? In particular, does a strong boundary block PFC?

(b) Is boundary strength marked by both F_0 and duration? How are boundary strength and focus encoded simultaneously in intonation?

(c) Does newness of postfocus words have any effect on PFC?

2 Methods

2.1 Stimuli

The key to our experimental design was to make sure that the factors under examination were controlled not by their prosodic patterns, but by nonprosodic factors, so as to avoid circularity in data interpretation. The three factors controlled in the study were boundary strength, focus, and newness of postboundary words. Boundary strength was manipulated by varying syntactic structures. Focus was controlled by context sentences that induce emphasis on different words in the subsequent target sentences. Four focus conditions were included: focus right before the boundary (focus on word X), focus right after the boundary (focus on word Y), final focus (focus on word Z), and neutral focus. Newness of the postboundary part was controlled by presence or absence of the postboundary words in the preceding context sentence. Here, postboundary new means that no part of the constituent after the boundary is mentioned in the previous context. In this way, the three factors were manipulated independently.

Another motivation of the experiment design was to set aside theoretical controversy on boundary categories, which is unsettled due to conceptual ambiguities as discussed in the Introduction. In the current study we asked a simpler question: do different degrees of boundaries have different acoustic cues? We chose two levels of boundary strengths, weak, medium, and strong, without assuming that they are of particular categories. These relative boundary strengths were constructed based on syntax, referred to as B1 (weak), B2 (medium), and B3 (strong) boundaries, respectively. The examples of the target sentences are listed below in sentences 1a–1c. B1 boundary was weak because it was within a compound word. B3 boundary was strong because it separated two coordinate clauses and was marked by a comma in text. B2 boundary was in-between because it was a juncture between the subject noun phrase and the verb phrase of a sentence, in which the subject noun phrase is a relative clause. In the case of the strong boundary (B3), speakers would naturally pause because of the punctuation. But a pause was less expected in the case of the B2 boundary and was not expected in the case of the B1 boundary.

(1) Example base sentences:

(1a) The first base sentence with B1 boundary:
我买的[柚栗]$_X$[兜]$_Y$送给[毛奶奶]$_Z$了。
wo3 mai3 de [you4li4] $_X$ [*dou1*] $_Y$ song4gei3[mao2nai3nai] $_Z$ le.
I buy NOM nut *tote bag* give to Maonainai ASP.
The "nut" tote bag I bought was given to Maonainai.

(1b) The first base sentence with B2 boundary:
我买的[柚栗]$_X$[都]$_Y$送给[毛奶奶]$_Z$了。
wo3 mai3 de [you4li4] $_X$ [*dou1*] $_Y$ song4gei3[mao2nai3nai] $_Z$ le.
I buy NOM nut *all* give to Maonainai ASP.
The nuts that I bought were all given to Maonainai.

(1c) The first base sentence with B3 boundary:
我买了[柚栗]$_X$, [都]$_Y$送给[毛奶奶]$_Z$了。
wo3 mai3 le [you4li4] $_X$, [*dou1*] $_Y$ song4gei3[mao2nai3nai] $_Z$ le.
I buy ASP nut, all give to Maonainai ASP.
I bought nuts and gave them all to Maonainai.

Note that our strategy here is to use monosyllabic homophones to form different syntactic structures. This is to make sure that sentences in different boundary conditions are as similar as possible. In condition B1, the boundary between word X (you4li4, the name of a special chestnut) and Y (dou1, "tote bag") is the weakest, as X is a modifier of Y and XY is a compound noun [you4li4dou1, Youli tote, meaning a tote with this kind of special chestnut painted on it]. In condition B2, a homophone of Y (dou1, written as a different character) with the meaning "all" is used so that the boundary between

DOI: 10.1159/000453082

X and Y is stronger than in B1, as it is a juncture between the subject (word X) and the verb (word Y). Word X here is the head of the relative clause. In condition B3, an aspect marker “le” replaces the nominal marker “de” in the first part of the sentence so that the preboundary part forms a matrix clause, and the boundary between X and Y is even stronger, as it is a juncture between two clauses marked by a comma. In this way, three boundary strengths are clearly distinguished, with no need for explanation to the speakers. The phonemes are mostly identical across the three boundary conditions except that the preboundary function word in B3 differs from that in B2 and B1. Moreover, both function words are in the neutral tone and in the same position of the sentence so that the difference of intonation caused by them should be limited. In the second set of sentences, word X is “lvwa” (green frog) (see the Appendix). Here, the main reason for selecting “Youli” (a made-up name for some kind of nuts) and “Lvwa”(green frog) is to use as many sonorant consonants as possible so as to obtain continual F_0 in the target words.

With this design we could be sure that any difference found between the sentences would come from variations in boundary strength rather than due to lexical, tonal, topical, or other factors. As B2 and B1 sentences were phonemically identical and very similar in sentence structure, we also expected that the difference between these two boundaries would not be as large as the difference of either of them from B3.

Focus and newness were easy to manipulate while keeping the target sentences the same across the conditions. This was done by using a preceding context sentence to induce either a corrective or information focus in the target sentence; see examples in sentences 2a and 2b below. Four focus conditions were constructed by putting focus on word X, Y, Z, or none of the words (neutral focus), in which the manipulated boundary was between X and Y. In this manipulation, the postboundary word Y was either in focus (focus on Y), prefocus (focus on Z), postfocus (focus on X) or neutral focus.

To control newness of the postboundary words, we could either include (given) or exclude (new) them in the preceding context sentence, as shown in sentences 2a and 2b below. Also with this manipulation, when postboundary words were new, the focus was corrective, whereas when the postboundary words were given, the focus was informational. However, from previous studies (Baumann et al., 2007; Hanssen et al., 2008; He et al., 2011; House and Sityaev, 2003; Hwang, 2012; Katz and Selkirk, 2011; Kügler and Ganzel, 2014; Sahkai et al., 2013), no systematic acoustic difference was found between information and corrective focus. Thus, the differences between the given and new conditions, if any were found, would be evidence that newness is encoded in intonation. On the other hand, if no difference between given and new conditions is found, it is more likely that neither newness nor focus type leads to systematic intonational variation. It is unlikely that these two factors counterbalance their effects in intonation, because corrective focus (also the new condition) presumably has stronger prosodic marking than information focus (also the given condition). Finally, neutral focus was associated with only the new condition, because it would be unnatural to have a whole target sentence repeated twice to make a given condition. Neutral focus is usually interpreted as an answer to a question such as “what happened.” Then, it is not possible to put the target sentence in the question. As will be seen in the Results section, two steps of data analysis are taken to deal with the situation that there is no neutral-given condition.

(2) Preceding context for creating 4 focus conditions and 2 conditions of newness for the postboundary content.

Here, we take the sentence with B3 boundary as an example and give only the Chinese characters and English translations in order to reduce clutter. Words that are underscored (Chinese) or in capital letters (English) are in focus:

(2a) Postboundary given:

Focus on X: 你是问，我买了什么，都送给毛奶奶了？我买了**柚栗**，都送给毛奶奶了。

Are you asking what I bought and gave them all to grandma Mao? I bought NUTS and gave them all to Maonainai.

Focus on Y: 你是问，我买了柚栗，送没送给毛奶奶？我买了柚栗，**都**送给毛奶奶了。

Are you asking I bought nuts, did I give them to Maonainai or not? I bought nuts and gave them ALL to Maonainai.

Focus on Z:

你是问，我买了柚栗，都送给谁了？我买了柚栗，都送给**毛奶奶**了。

Are you asking I bought nuts, who did I give them to? I bought nuts and gave them all to MAONAINAI.

(2b) Postboundary new:
Focus on X:
不是甜橙。我买了**柚栗**，都送给毛奶奶了。
It is not orange. I bought NUTS and gave them all to Maonainai.
Focus on Y:
不是一半。我买了柚栗，**都**送给毛奶奶了。
It's not half. I bought nuts and gave them ALL to Maonainai.
Focus on Z:
不是李妈妈。我买了柚栗，都送给**毛奶奶**了。
It's not Limama. I bought nuts and gave them all to MAONAINAI.
Neutral focus:
我要告诉你一件事。我买了柚栗，都送给毛奶奶了。
I need to tell you something. I bought nuts and gave them all to Maonainai.

Two sets of basic sentences were constructed with different tones, words, and homophones of Y (*dou1* meaning either "tote" or "all" in the first set of sentences and *hui4* meaning either "stew" or "will" in the second set of sentences). The tones of word X in the first and the second sentence sets were FF and FH, respectively. In the FF tone sequence, the falling tone is in the same direction of postfocus F_0 lowering, whereas in the FH sequence the high tone is in the opposite direction. In data analysis, we will take these into consideration.

There were thus 2 (newness) × 3 (focus) × 3 (boundary) × 2 (base sentence) = 36 sentences with focus on either X, Y, or Z words, and 2 (base sentence) × 3 (boundary condition) = 6 neutral sentences. In total, there were 42 unique sentences. For a full list of reading materials, see the Appendix. Eight speakers each recorded 3 repetitions of these sentences, thus providing 42 × 3 × 8 = 1,008 sentences for analysis.

2.2 Predictions

Based on the literature review above, we made the following predictions for the three research questions.

(a) Does PFC apply across boundaries of different strengths? In particular, would a strong boundary block PFC?

PFC is expected in B1, and possibly also in B2. But the predictions are open as to whether PFC applies in B3. It is possible that the B3 boundary will block PFC.

(b) Is boundary strength marked by both F_0 and duration? How are boundary strength and focus encoded simultaneously in intonation?

Durational differences across the three boundary conditions are expected, with silent pauses occurring in the B3 condition, and possibly in the B2 condition, but not in the B1 condition. The amount of preboundary lengthening may also vary across the three boundaries. F_0 variations, such as preboundary lowering and phrase-initial rising, are probably larger across bigger boundaries. There might be some boundary-initial lengthening at a stronger boundary as well.

For the interaction between boundary strength and focus, there are at least two possibilities, i.e., one function may overtake the other one, or they are mostly independent of each other. In the first case, when a boundary-final word is in focus, it may not be lengthened further due to focus since the effect of final lengthening already applies. In the second case, final lengthening and lengthening due to focus may apply at the same time. For F_0 variation, the effect of final lowering and in-focus F_0 raising could cancel each other.

(c) Does newness of postboundary words have any effect on PFC?

This will depend on (i) whether newness is cued by F_0 at all and (ii) if it is, whether the F_0 variation due to newness is also detectable postfocally. Based on our previous study (Wang and Xu, 2011), we predict that there is no systematic effect of newness on postfocus F_0.

2.3 Speakers

Eight speakers, 5 females and 3 males aged between 20 and 25 years, participated in the experiment. They were all born and brought up in Beijing, and spoke Beijing Mandarin as their native language without speaking other dialects. They did not report any speech and hearing impairments. They were paid a small amount of money for their participation.

2.4 Recording Procedure

The subjects were recorded individually in the speech lab at Minzu University of China. They were asked to read aloud both the context sentences and target sentences at a normal speed and in a natural way. They sat before a computer monitor, on which the test sentences were displayed, with the focused words highlighted, using AudiRec, a custom-written recording tool. The highlighting of the focused word was to make the reading task a little easier for the speakers, because the location of the focused word was already contextually determined. The subjects were asked to read both the context sentence and target sentence naturally, paying attention to both the text and punctuations. The speakers were asked to go through all the target sentences and read them in silence before the recording. During the recording, when the experimenter determined that a particular sentence was not said properly, e.g., with wrong pronunciation or disfluency, the subject was asked to say the whole discourse again. This happened only occasionally. A Shure 58 Microphone was placed about 10–15 cm in front of the speaker. All sentences were digitized directly into a Thinkpad computer and saved as WAV files. The sampling rate was 48 kHz, and the sampling format was 1 channel 16-bit linear. Each subject read the sentences three times, once in each session, with about 5-min breaks between sessions. In each session, all the 42 sentences were randomized, and each subject had a different randomization order. The total recording time was less than 1 h, with a 5-min practice at the beginning.

2.5 Acoustic Measurement

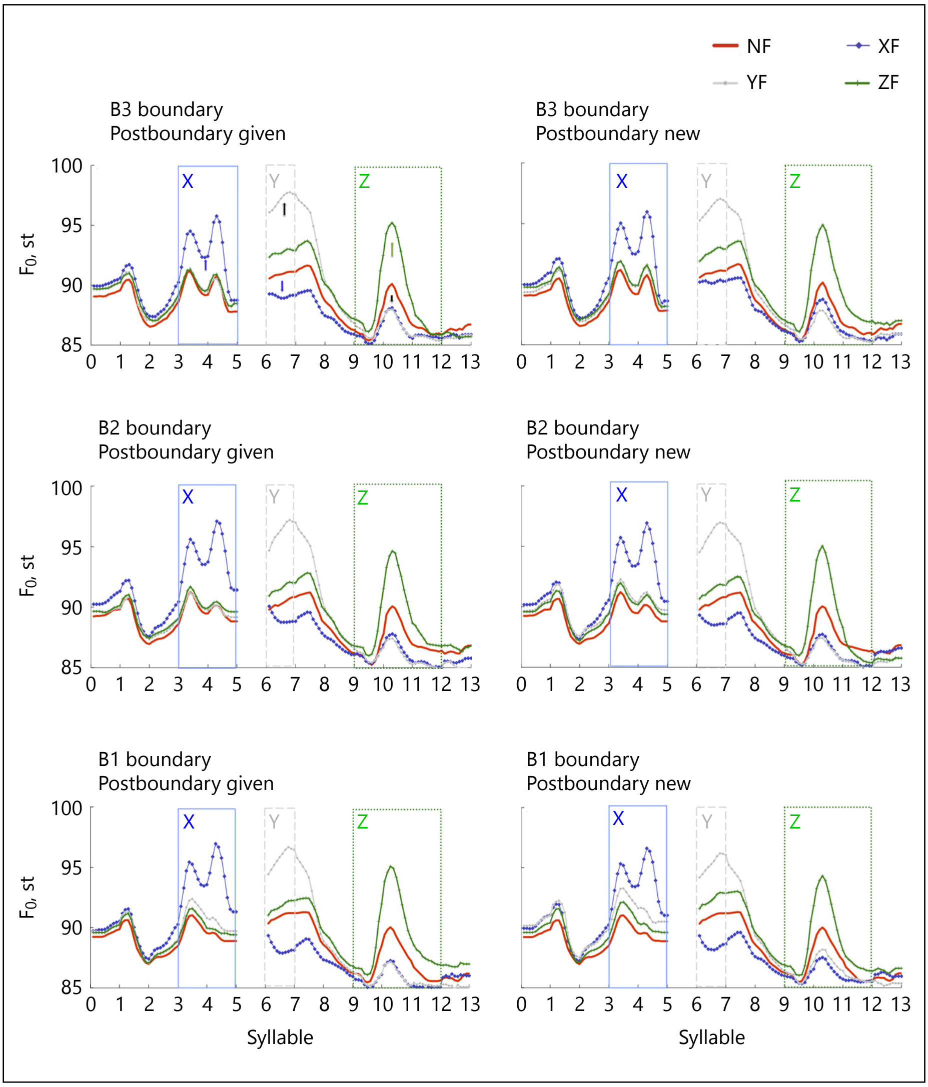
The target sentences were extracted and saved as separate WAV files. The acoustic analysis procedures were similar to those in Wang and Xu (2011). ProsodyPro, a Praat script (Xu, 2013) running under Praat (Boersma and Weenink, 2005), was used to take F_0 and duration measurements from the target sentences. With ProsodyPro, the first and the third authors segmented the target sentences into syllables, and at the same time hand-checked vocal cycle markings generated by Prosody for errors, such as double marking and period skipping. ProsodyPro then generated syllable-by-syllable F_0 contours that are either time normalized or in the original time scale. At the same time, the script extracted various measurements, including maximum F_0 and minimum F_0 of the four target words (word X, Y, Y + 1 and Z).

3 Results

In the following presentation of the results, we will first compare time-normalized F_0 contours. The time normalization enables averaging across repetitions as well as speakers, which makes it possible to directly compare contours in fine detail. The graphical analysis is then followed by quantitative analyses, in which all the measurements, including duration, maximum F_0, and minimum F_0, were taken from F_0 contours on the original time scale.

3.1 Graphical Comparison of F_0 Contours

In order to show the effects of the key variables clearly, we will first compare two of the variables, while keeping the third variable constant. Figure 1 displays time-normalized F_0 contours of sentence set 1 in the three boundary and two newness conditions, with the 4 focus conditions overlaid in each plot. All the contours were averages across 8 speakers and their 3 repetitions. We can see that all focused words show raised F_0 relative to the neutral focus F_0 (as indicated by the upward arrows). This shows that in-focus raising of F_0 occurred in all the boundary and newness conditions. F_0 of post-focus words is lowered in all the boundary and newness conditions (as indicated by the downward arrows) relative to the corresponding neutral focus words in both X and Y focus conditions. The most interesting contours are the sentences in the X focus condition. They show clearly that a strong prosodic boundary (B3) does not block postfocus F_0 compression.



Color version available online

Fig. 1. Mean time-normalized F_0 contours of sentence set 1 in different boundary conditions (from top to bottom: B3, B2, B1) and in two newness conditions (left column: the postboundary words are given; right column: the postboundary words are new) with four focus conditions overlaid in each graph. Boundary strength variations occur between syllables 5 and 6, as indicated by the gap in the F_0 contours. The NF (neutral focus) condition is under the condition that postboundary words are new.

The plots of sentence set 2 (Fig. 2) show very similar patterns, except that when word X is in focus; PFC is clearly seen in the last word, but not in the words right after word X (i.e., word Y) in B1 and B2 conditions. However, in the B3 condition, PFC *can* be seen in word Y. This seems to be due to a familiar *carryover* effect (Chen and Xu, 2006; Wang and Xu, 2011; Xu, 1999). That is, as introduced in the Material section, the final in-focus syllable in sentence set 2 has a high tone (contrary to the falling tone in sentence

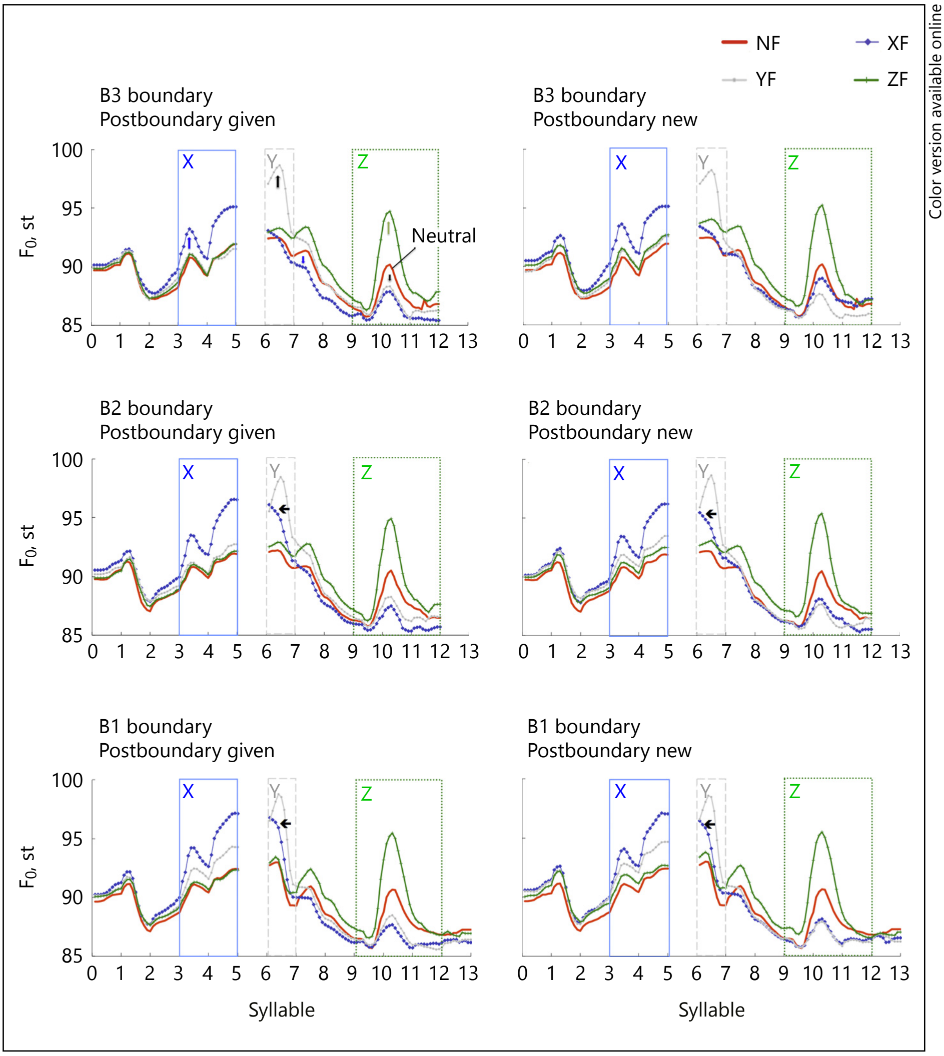


Fig. 2. Mean time-normalized F_0 contours of sentence set 2 in different boundary conditions (from top to bottom: B3, B2, B1) and in 2 newness conditions (left column: the postboundary words are given; right column: the postboundary words are new) with 4 focus conditions overlaid in each graph. Boundary strength variations occur between syllables 5 and 6, as indicated by the gap in the F_0 contours. The NF (neutral focus) condition is under the condition that postboundary words are new.

set 1), which, when exaggerated by focus, generates a rising momentum in the opposite direction as the F_0 PFC. This rising momentum takes time to reverse by PFC. Such a reversal time seems available in the B3 condition thanks to the pause (which, as shown later, is over 200 ms), but unavailable in B1 and B2 due to lack of pause. Regardless of this difference, F_0 PFC seems to apply across boundaries of all three strengths.

Next, to check whether the amount of PFC differs across the boundary conditions, we put sentences with different boundary strengths together in Figure 3. To save space,

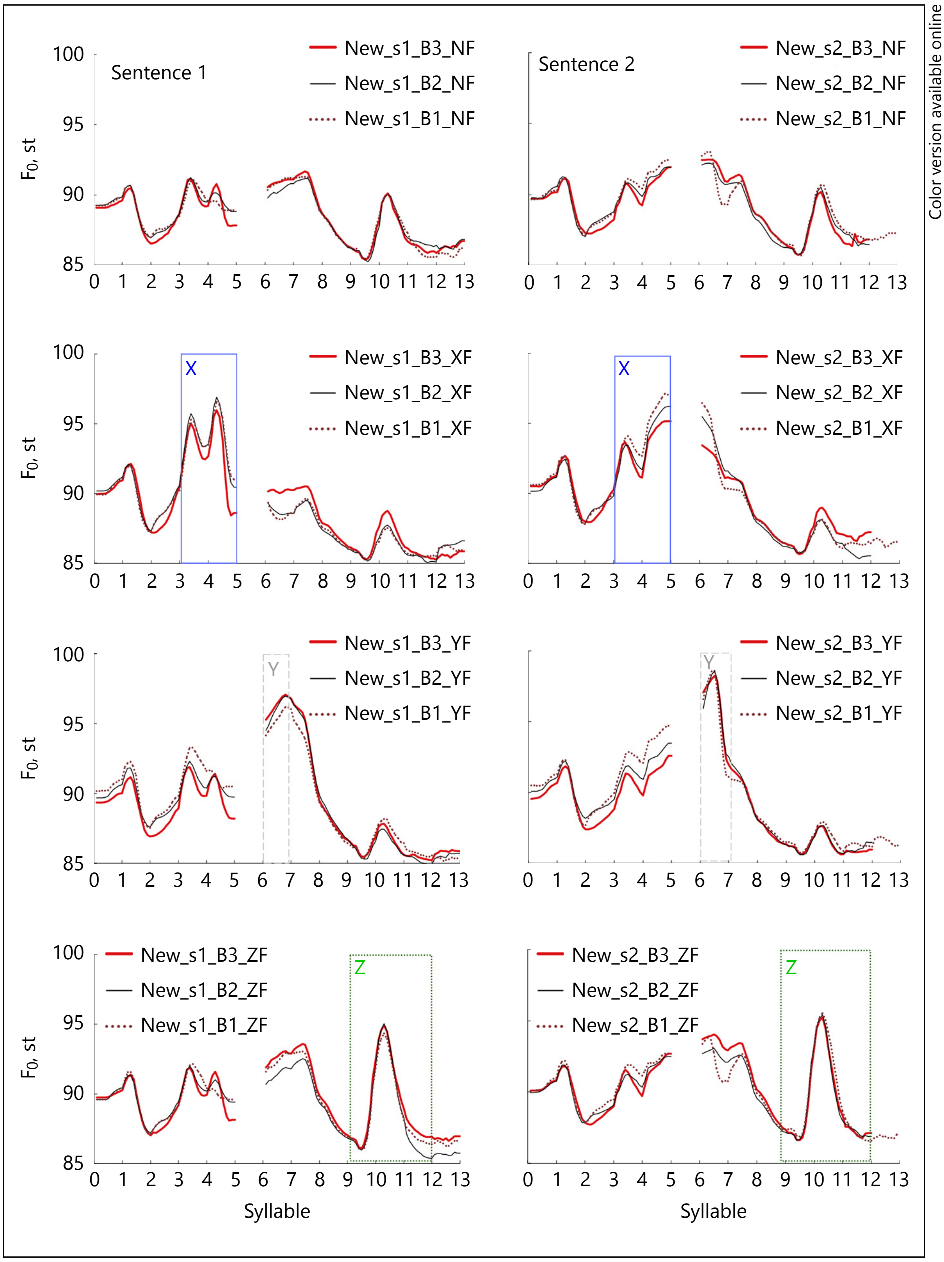


Fig. 3. Mean time-normalized F_0 contours of sentence set 1 (left) and sentence set 2 (right) in all focus conditions with three boundary conditions overlaid in each graph. From top to bottom are neutral focus (NF), focus on X, Y, and Z word separately. In all these sentences the postboundary words are new. The B1 contours in sentence set 2 (right column) are 1 syllable longer because of the sentence final particle "le."

here we only present sentences in the new condition. In Figure 3, we can see that preboundary syllables (word X) tend to have a lower F_0 in B3 than in the other two boundary conditions, whereas the difference between B1 and B2 conditions is relatively small. This is true of all the focus conditions. In the postboundary syllable (word Y), sentences with B3 boundary do not show systematically higher F_0 than those with B1 and B2 boundaries. Sentences with B1 and B2 boundaries do not seem to differ much in any of the focus conditions. And the nonboundary words, i.e. those either before word X or after word Y, do not show any systematic variation of F_0 across different boundary conditions either. Sentences in the *given* condition show roughly the same pattern.

Finally, Figure 4 compares sentences in the given and new conditions. Shown here are the F_0 contours of sentence 1 in the three boundary and three focus conditions. In general, there is not much difference between the given and new conditions in any of the boundary and focus conditions. The intonation contours of sentence set 2 (not shown here in the interest of space) show a similar pattern.

In summary, from the graphical comparisons, we can see that focus has a systematic effect on F_0: in-focus F_0 raising and postfocus F_0 lowering occur in all boundary and newness conditions. The F_0 variation due to boundary strength shows visible yet small differences around the boundaries, which are mostly in terms of lowering the preboundary word under the B3 boundary condition. The newness of the postboundary words does not show any clear impact on F_0.

3.2 Quantitative Analysis of F_0

The main goal of the quantitative F_0 analysis is to test the significance of the effects of focus, boundary and newness and their interactions. This is achieved by 3-way repeated-measures ANOVAs on the maximum and minimum F_0 of the four target words (word X, Y, Y + 1 and Z) in the two sentence sets. Because the neutral focus sentences only have the postboundary words being new, but not given (see section 2.1), they are not included in the 3-way ANOVAs. The 2-way interaction of boundary and newness in the neutral focus condition will be analyzed later in section 3.4. In addition, because the focus effect is largely known, it is not critical for the current study to compare other focus conditions with neutral focus condition. Besides, because prefocus F_0 is known to remain largely unchanged from the neutral focus F_0 (Chen and Gussenhoven, 2008; Xu, 1999), the final-focus condition (Z focus) can serve as the baseline for assessing the size of postfocus F_0 lowering.

3.2.1 Maximum F_0

Table 1 shows the maximum F_0 of the four target words (X, Y, Y + 1 and Z) in sentence sets 1 and 2, in semitones (st; calculated by 12 $\log_2[F_0]$). The maximum F_0 of the word after the Y word (Y + 1) is also shown here because, as seen in the graphical analysis, there is an effect of carryover articulatory velocity in the X focus condition of sentence set 2, so the PFC effect of the X focus is manifested mainly in the Y + 1 word. The results of 3-way repeated-measures ANOVAs on the maximum F_0 of the three target words in the two sentence sets are shown in Table 2. The 3 independent variables are focus (focus on X, Y, or Z word), boundary (B3, B2, B1) and newness (postboundary words being given or new).

We can see in Table 2 that the focus effect is significant in all the words and in both sentences. In Table 1, focused words have the highest maximum F_0, whereas postfocus words have the lowest maximum F_0. On average, maximum F_0 of in-focus, postfocus

Color version available online

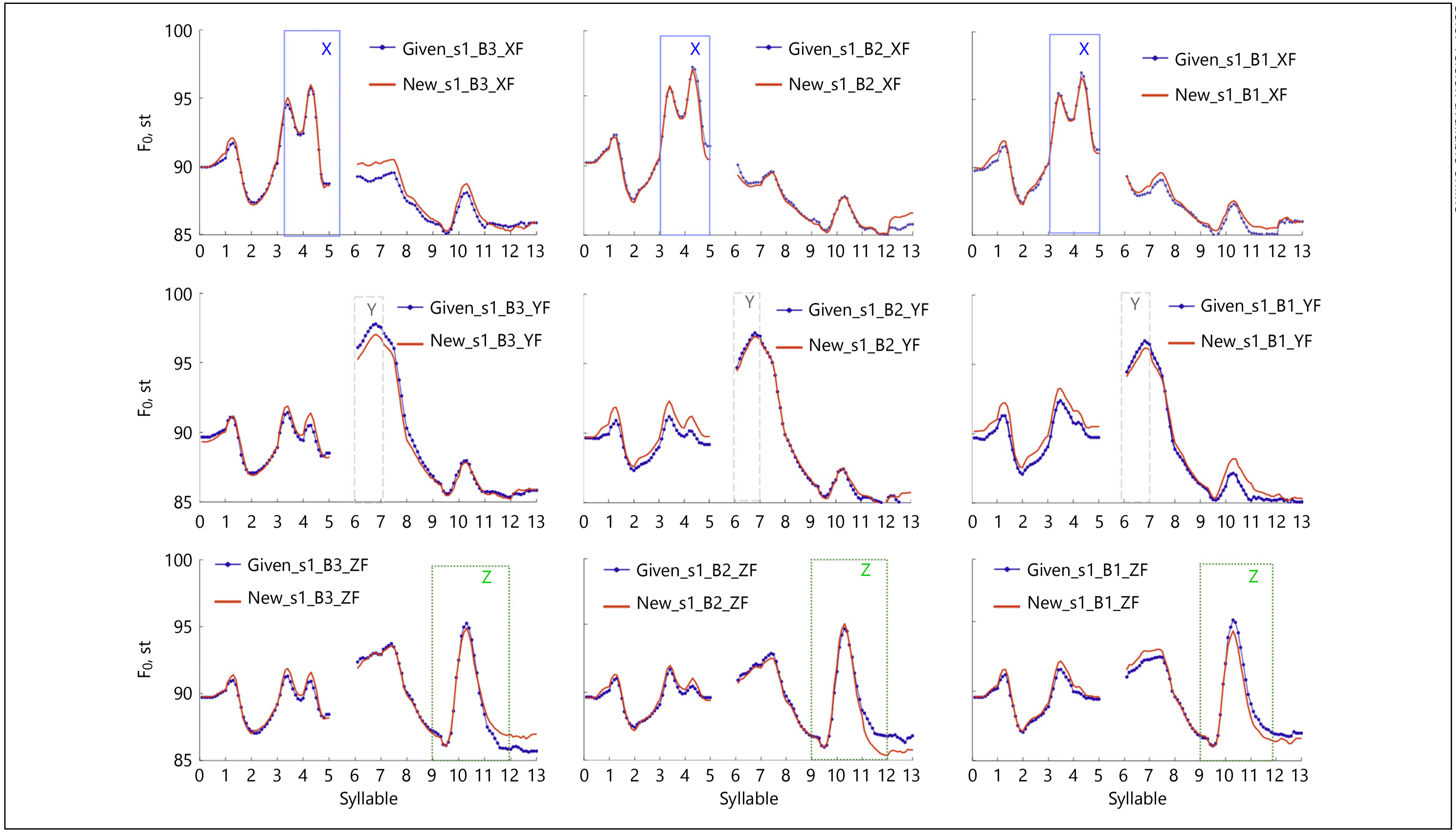


Fig. 4. Mean time-normalized F_0 contours of sentence set 1 comparing 2 newness conditions in the 3 boundary conditions (from left to right: B1, B2, B3) and the 3 focus conditions (from top to bottom: X, Y and Z focus).

Table 1. Maximum F_0 (st) of the 4 target words in different focus and boundary conditions when the postboundary part is given

	B1 boundary			B2 boundary			B3 boundary		
	XF	YF	ZF	XF	YF	ZF	XF	YF	ZF
Sentence set 1									
Word X	***97.2***	92.8	90.7	***97.3***	92.8	91.0	***96.2***	91.2	91.3
Word Y	90.0	***96.8***	92.5	90.4	***97.4***	92.3	89.6	***97.9***	93.3
Word Y + 1	89.2	96.0	92.9	89.7	96.7	93.1	89.7	97.5	94.0
Word Z	87.4	87.5	***95.3***	87.9	87.6	***95.2***	88.5	88.2	***95.5***
Sentence set 2									
Word X	***96.9***	95.2	92.3	***96.4***	93.9	92.3	***95.2***	92.4	92.1
Word Y	96.6	***98.9***	93.2	95.7	***98.5***	93.0	92.7	***98.8***	93.2
Word Y + 1	89.8	90.9	92.2	90.6	92.2	92.7	89.8	92.2	93.4
Word Z	87.5	87.4	***95.3***	87.2	87.7	***95.2***	87.7	87.7	***94.9***

Figures in italics indicate values of focused words.

Table 2. Results of 3-way repeated-measures ANOVA on maximum F_0 of the 4 target words

Sentence	Word	Focus F(1, 7)	Boundary F(2, 14)	Givenness F(2, 14)	Interaction
Set 1	X	225.398***	7.798*	5.496*	B × F: 6.015*
	Y	256.06***	4.603, ns	6.015, ns	–
	Y + 1	119.4***	11.351***	0.139, ns	–
	Z	313.925***	5.461*	0.220, ns	–
Set 2	X	64.519***	22.737***	0.117, ns	–
	Y	109.248***	12.078**	0.638, ns	B × F: 26.967***
	Y + 1	16.528***	16.752***	0.018, ns	B × F: 5.205**
	Z	199.351***	1.262, ns	3.405, ns	–

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. For the interaction, B and F stand for boundary and focus, respectively.

and prefocus words are 96.6, 89.3, and 92.5 st, respectively. Of the most interest to us is that in B3 sentences, PFC also applies. Looking at the maximum F_0 values of B3 sentences (the last column of Table 1), if taking Z focus as a baseline, word Y and Y + 1 under the X focus condition are lowered by 2.1 and 4 st on average.

For the effect of boundary strength on maximum F_0, we can see in Table 2 that the preboundary word (word X) shows a significant effect in both sentences. Figure 5 compares maximum F_0 of word X under the three boundary conditions when the focused word is either X, Y, or Z. We can see that maximum F_0 in B3 is lower than in B1 and B2 sentences, and this holds when focus is on word X and word Y, but not in the Z focus condition. This effect seems to be due to a phrase-final F_0 lowering, which is related to the well-known sentence-final lowering (Liberman and Pierrehumbert, 1984). The absence of this boundary effect in Z focus is interesting, as it suggests that the anticipation of an

Table 3. Minimum F_0 (st) of the 4 target words in different focus and boundary conditions when the postboundary words are given

	B3 boundary			B2 boundary			B1 boundary		
	XF	YF	ZF	XF	YF	ZF	XF	YF	ZF
Sentence set 1									
Word X	***89.0***	88.3	88.4	***90.4***	89.5	89.4	***90.4***	90.2	89.5
Word Y	89.3	***95.5***	92.1	88.4	***93.8***	90.7	88.1	***93.6***	90.9
Word X	***91.0***	89.5	89.4	***91.8***	90.5	90.0	***92.6***	91.7	90.6
Word Y	90.2	***92.4***	92.3	91.1	***92.4***	91.4	89.9	***91.1***	90.2

Figures in italics indicate values of focused words.

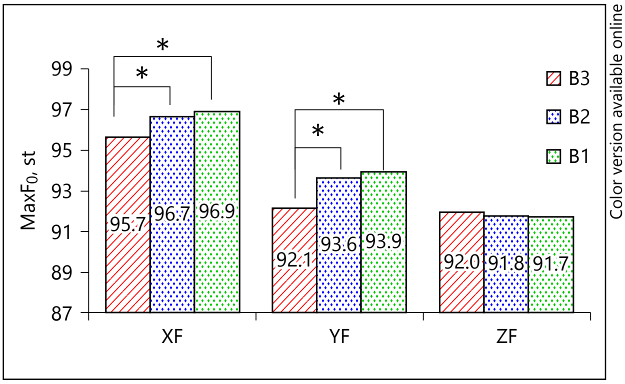


Fig. 5. Mean maximum F_0 of word X in different boundary and focus conditions.

upcoming focus at the end of the whole sentence somehow forces the two constituent phrases more closely together.

For the boundary effect in the Y word, Table 2 shows that it is significant in sentence set 2 but not in sentence set 1. That is, the maximum F_0 of the postboundary word (word Y) is lower in B3 than in B1 and B2 in the X focus condition. As discussed earlier, the articulation time provided by the long B3 boundary is likely sufficient for PFC to reverse the rising momentum of the preceding in-focus high tone. Given its articulatory nature, this effect is largely a byproduct of tonal articulation relating to focus rather than a genuine boundary effect.

Finally, Table 2 also shows that the difference between the two newness conditions is not significant in most cases. Only the X word in sentence set 1 shows an effect at the $p < 0.05$ level (93.4 vs. 94.1 st in given and new conditions on average). Newness of postboundary words does not seem to have any direct impact on F_0.

3.2.2 Minimum F_0

Table 3 presents minimum F_0 of word X and Y in different focus and boundary conditions when postboundary words were given. The results of minimum F_0 under the new condition are similar to what is shown in Table 3. Table 4 shows the results

Table 4. Results of 3-way repeated-measures ANOVA on minimum F_0 of the four target words

Sentence	Word	Focus F(1, 7)	Boundary F(2, 14)	Givenness F(2, 14)	Interaction
Set 1	X	1.901, ns	6.957*	0.543, ns	–
	Y	226.112***	12.195**	0.121, ns	B × F × N: 4.982***
	Y + 1	20.764***	2.34, ns	0.400, ns	–
	Z	4.603*	1.387, ns	0.432, ns	–
Set 2	X	14.449***	5.261*	1.222, ns	–
	Y	6.687*	14.037**	1.313, ns	B × F: 9.909**
	Y + 1	27.816***	1.485, ns	1.639, ns	–
	Z	4.94*	2.222, ns	1.668, ns	–

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. For the interaction, B, F, and N stand for boundary, focus, and newness, respectively.

of 3-way repeated-measures ANOVAs on minimum F_0 of the four target words (X, Y, Y + 1 and Z), with focus, boundary, and newness as independent variables. We can see in Table 4 that there were boundary effects only in word X and word Y. In the B3 condition, the preboundary word ends with a lower minimum F_0 (89.2, 90.3, and 90.9 st in B3, B2, and B1 conditions, respectively, averaged across all focus conditions) and the postboundary word starts with a higher minimum F_0 (92.0, 91.3, and 90.6 st in B3, B2, and B1 conditions, respectively, averaged across all focus conditions). Post hoc tests show that preboundary lowering and postboundary raising in minimum F_0 reach the significant level of $p < 0.05$ only between B3 and B1 conditions, whereas the B2 boundary shows no significant difference from either B1 or B3. Moreover, these two boundary effects hold for all the focus conditions (Table 4).

As for focus effect, it is significant in almost all the words (Table 4), except for word X in the first sentence. The values in Table 3 show that focused words have higher minimum F_0 than their unfocused counterparts. On average, minimum F_0 of word X in the X focus condition is 1.1 st higher than that in the Y focus and Z focus conditions.

Finally, we can see in Table 4 that newness does not have effects on any of the words. Neither is there any interaction between newness and boundary or between newness and focus.

Overall, the results of the graphic and quantitative analyses on F_0 can be summarized as follows. (1) Focus has a stable trizone realization in all the newness and boundary conditions: prefocus F_0 is largely intact, in-focus F_0 is raised, and postfocus F_0 is lowered, in terms of both maximum and minimum F_0. Most importantly, PFC applies across a strong boundary. (2) A boundary effect on F_0 mostly occurs in the preboundary word, with lower ending F_0 before a stronger boundary, in terms of both maximum and minimum F_0. A postboundary word seems to start with higher minimum F_0 when the boundary is stronger but shows no difference in maximum F_0. However, the boundary effect in F_0 is not equally sensitive to all the three boundaries. Most of the boundary difference is seen between B3 and B1/B2 conditions, with no clear difference between B1 and B2 conditions. Moreover, preboundary F_0 lowering is absent when focus is toward the end of a sentence. (3) Newness of the postboundary part does not show any clear effect in either maximum or minimum F_0.

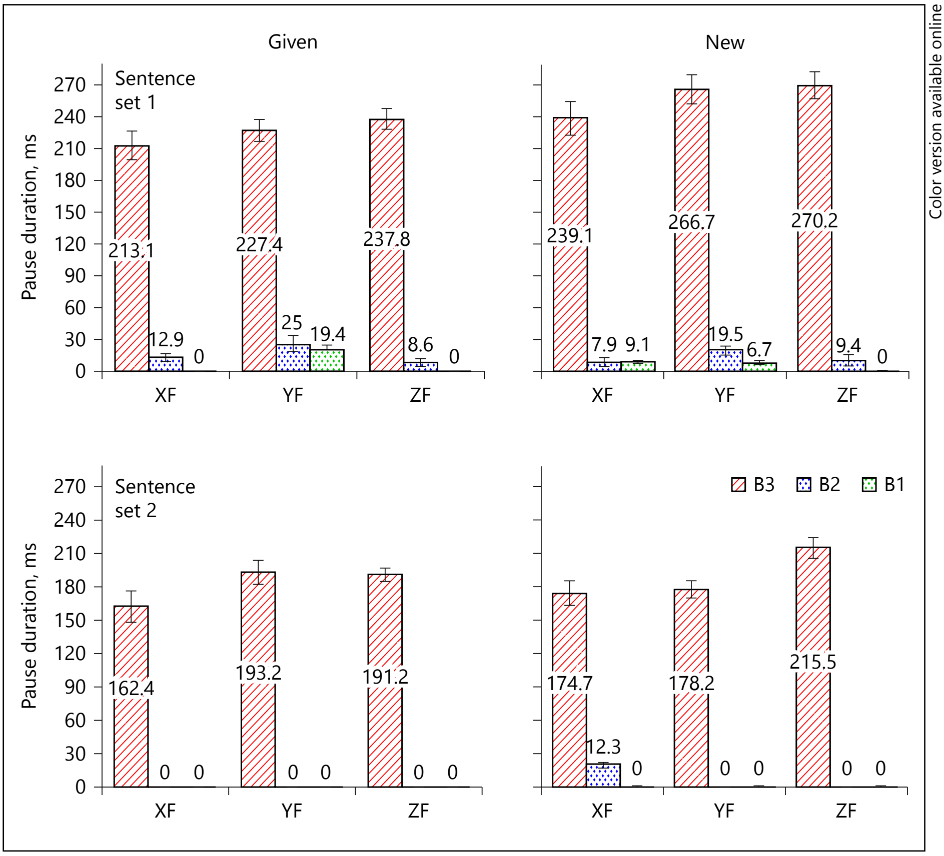


Fig. 6. Pause duration of the 2 sentence sets in different focus, boundary and newness conditions. The vertical bars represent standard errors.

3.3 Duration

3.3.1 Silent Pause

Figure 6 displays duration of silent pause in the two sentence sets in different focus, boundary and newness conditions. B3 has much longer silent pauses than the other 2 boundaries in all the focus and newness conditions (214.1 ms, 8.0 ms and 2.9 ms for B3, B2, and B1 conditions, respectively). Here in B1 and B2 conditions, silent pauses occur only occasionally. Two 3-way repeated-measures ANOVAs, with focus, boundary, and newness as independent variables, show that boundary has an effect (sentence set 1: $F(2, 14) = 174.851$, $p < 0.001$; sentence set 2: $F(2, 14) = 264.005$, $p < 0.001$), but focus does not (sentence set 1: $F(2, 14) = 1.351$, ns; sentence set 2: $F(2, 14) = 0.835$, ns). Newness has a marginal effect on silent pause in sentence set 1 ($F(1, 7) = 6.034$, $p = 0.044$) but not in sentence set 2 ($F(1, 7) = 1.216$, ns). In sentence set 1, when the postboundary words are new, pause duration is slightly longer than when they are given (258 vs. 226 ms). However, this effect is not stable, as it disappears in sentence set 2.

To further test whether focus location and newness have any effect on pause duration, we compared B3 boundary in a 3-way repeated-measures ANOVA, with sentence,

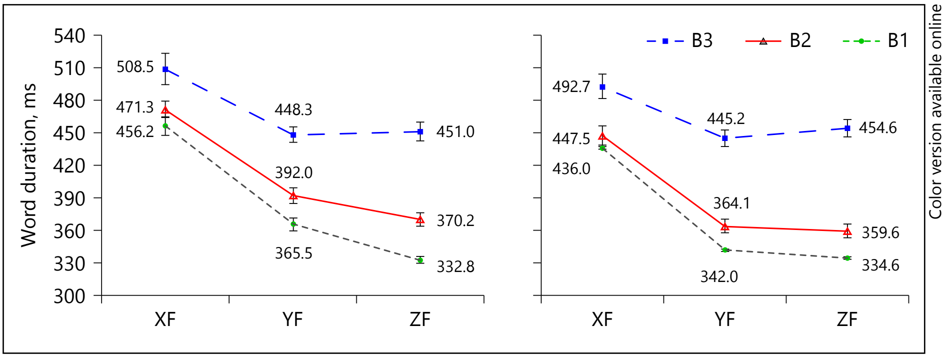


Fig. 7. Duration of word X in different boundary and focus conditions (left: sentence set 1; right: sentence set 2), averaged across 2 newness conditions.

newness, and focus as independent variables. Both sentence (F(1, 7) = 32.226, $p < 0.001$) and newness (F(1, 7) = 6.889, $p = 0.034$) have main effects, but not focus (F(2, 14) = 1.317, ns). At a B3 boundary, the pause duration is 242.4 and 185.9 ms in sentence sets 1 and 2, respectively. Focus does not change pause duration. However, pause duration is slightly longer when the postboundary part is new than when it is given, which mainly comes from sentence set 1 as discussed above.

3.3.2 Phrase-Final Lengthening

Figure 7 displays durations of word X in the three focus conditions in three boundary conditions. The corresponding values of the two newness conditions are averaged (as we will see later that newness has no effect). We can see that the duration of the preboundary word increases as boundary gets stronger in all the focus conditions, and the difference between B3 and the other two boundaries is greater than the difference between B1 and B2 conditions. Also, when word X is focused, it has the longest duration.

We applied 3-way repeated-measures ANOVAs on phrase-final duration for the two sentence sets separately, with focus, boundary, and newness as independent variables. Results show that boundary (sentence set 1: F(2, 14) = 24.938, $p < 0.001$; sentence set 2: F(2, 14) = 32.87, $p < 0.001$) and focus (sentence set 1: focus: F(2, 14) = 58.719, $p < 0.001$; sentence set 2: F(2, 14) = 63.156, $p < 0.001$) both have significant main effects, and the interaction between them is also significant (sentence set 1: F(4, 28) = 8.612, $p = 0.006$; sentence set 2: F(4, 28) = 16.119, $p < 0.001$). Newness does not show any effect (sentence set 1: F(1, 7) = 4.915, ns; sentence set 2: F(1, 7) = 0.827, ns). Simple-effect tests show that there are significant differences among the 3 boundaries in Y focus and Z focus conditions. In the X focus condition, word duration of X in B3 boundary is significantly longer than in B1 and B2 boundaries, but there is no difference between B1 and B2 boundaries. This is true of both sentences.

From the above analysis, we can see that for the B3 boundary, not only phrase-final words lengthened, but also silent pause was inserted. Given that both preboundary lengthening and pause serve to signal a boundary, in Figure 8, the preboundary word duration and pause duration are combined, following Xu and Wang (2009). It can be seen that B3 stands out even more from the other two boundaries compared to Figure 7. B2 and B1 are different but in a much smaller scale. Similar 3-way repeated-measures

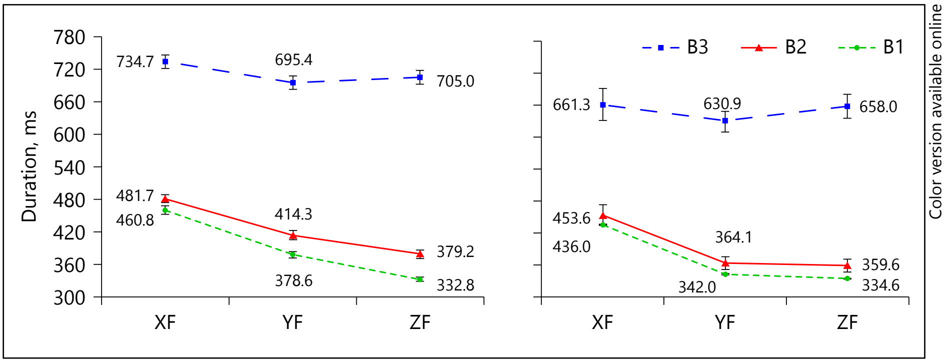


Fig. 8. The sum duration of word X and silent pause in different boundary and focus conditions (left: sentence set 1; right: sentence set 2), averaged across 2 newness conditions.

ANOVAs show effects of boundary (sentence set 1: F(2, 14) = 214.385, p < 0.001; sentence set 2: F(2, 14) = 235.086, p < 0.001), focus (sentence set 1: F(2, 14) = 46.752, p < 0.001; sentence set 2: F(2, 14) = 39.159, p < 0.001) and their interaction (sentence set 1: F(4, 28) = 14.449, p < 0.001; sentence set 2: F(4, 28) = 7.548, p = 0.009). Again, newness has an effect in sentence set 1 (F(1, 7) = 10.893, p = 0.013) but not in sentence set 2. Similar to the results of preboundary word duration, simple effect tests also show that the three boundaries differ significantly in the Y focus and Z focus conditions. In the X focus condition, the B3 boundary has the longest duration, but no difference between B1 and B2 boundary is found.

In general, a preboundary word shows a longer duration when it is in focus, in all the boundary conditions. As for boundary effect, a preboundary word shows a longer duration when the boundary is stronger, provided that the preboundary word is not focused. When the preboundary word is focused, a larger boundary still leads to longer preboundary word duration, but there is no difference between small boundaries.

3.3.3 Phrase-Initial Lengthening

The effect of phrase-initial lengthening is assessed with the duration of word Y, whose means in different focus and boundary conditions are displayed in Figure 9. First, we can see that word Y is the longest when it is in focus. In terms of boundary effect, word Y is much longer in B1 boundary than in the other two boundaries. This is because in B1, "dou1 (tote bag)" in the word "you4li4dou1 (Youli tote)" is actually the final syllable of a compound (as is also true of sentence set 2), whereas in B2 and B3 conditions, "dou (all)," though being a monomorphemic word, seems to have joined the following two syllables to form a trisyllabic phrase, which makes it phrase initial. Thus, the duration difference of word Y between the B1 and B2/B3 boundary is due to word structure rather than boundary degree. Between B2 and B3 conditions, while the word structure is the same, the duration of word Y is not much different (182.9 and 177.1 ms on average).

Separate 3-way repeated-measures ANOVAs were carried out for the two sentence sets, with newness, boundary, and focus as independent variables and duration of word Y as dependent variable. There are significant effects of boundary (sentence set 1: F(2, 14) = 66.075, p < 0.001; sentence set 2: F(2, 14) = 24.099, p < 0.001), focus (sentence

DOI: 10.1159/000453082

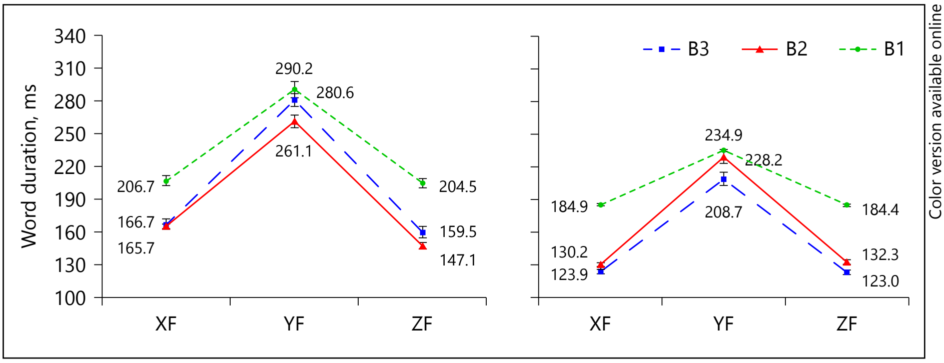


Fig. 9. Duration of word Y in different boundary and focus conditions (left: sentence set 1; right: sentence set 2), averaged across 2 newness conditions.

Table 5. Mean maximum F_0 of words X and Y in all focus and boundary conditions, with all sentences in the new condition

Word		Sentence set 1				Sentence set 2			
		NF	XF	YF	ZF	NF	XF	YF	ZF
X	B3	91.0	***96.4***	91.8	91.9	91.3	***94.9***	93.1	92.6
	B2	90.5	***97.3***	93.5	91.5	91.6	***95.8***	94.3	92.3
	B1	90.2	***96.8***	92.9	91.2	91.7	***96.6***	94.9	92.7
Y	B3	91.4	90.7	***97.3***	93.2	92.1	92.9	***98.2***	93.9
	B2	91.1	89.8	***97.1***	92.1	92.0	95.2	***98.4***	92.8
	B1	91.5	89.9	***96.3***	93.1	92.4	96.1	***98.2***	93.6

Figures in italics indicate values of focused words.

set 1: F(2, 14) = 54.751, $p < 0.001$; sentence set 2: F(2, 14) = 66.891, $p < 0.001$) and their interaction (sentence set 1: F(4, 28) = 17.102, $p = 0.021$; sentence set 2: F(4, 28) = 5.627, $p = 0.013$). Again, newness does not show any effect in either of the two sentence sets. Simple effect analysis shows that the Y word is significantly longer in B1 boundary than in the other two boundaries in all the focus conditions, and the difference between B2 and B3 conditions reaches significance level only in the Y focus condition, which is the case in both sentences. However, the two sentence sets show opposite directions, with the Y word being longer in B3 than in B2 in sentence set 1 (280.6 vs. 261.6 ms), but shorter in B3 than in B2 in sentence set 2 (208.7 vs. 228.2 ms). Overall, therefore, no clear systematic phrase-initial lengthening can be seen.

3.4 Further Analysis of Interaction between Boundary and Focus

In the preceding analyses, to investigate the interaction of focus, boundary, and newness, the neutral focus sentences have been excluded, because when the post-boundary part is given, there is no neutral focus condition. Here we include the neutral focus condition in a 2-way analysis to further examine the interaction between

Table 6. Results of 2-way repeated-measures ANOVA on maximum F_0 of words X and Y in all sentences in the new condition

Word	Sentence set 1			Sentence set 2		
	focus F(1, 7)	boundary F(2, 14)	interaction F(6, 42)	focus F(1, 7)	boundary F(2, 14)	interaction F(6, 42)
X	121.44***	1.704, ns	4.027*	55.204***	10.568**	5.546**
Y	113.919***	5.164*	–	87.269***	6.489*	13.699***

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

boundary and focus, with the newness factor excluded given its lack of effect on F_0. Table 5 presents mean maximum F_0 of word X and Y in the four focus and three boundary conditions when the postboundary part is new. Table 6 shows the results of 2-way repeated-measures ANOVAs.

Again, we can see that focus has effects on both words of the two sentence sets. In sentence set 1, word Y in the X focus condition has the lowest maximum F_0 in all the boundary conditions, due to PFC. It is not shown in sentence set 2, due to the carryover articulation explained before (see sections 3.1 and 3.2).

As for the boundary effect, consistent with the other focus conditions (Fig. 3), the B3 boundary shows a lower maximum F_0 in word X in the neutral focus condition. In word Y, it does not show any difference among the three boundaries in the neutral focus condition. These results are all consistent with the results shown in section 3.2.

4 General Discussion

This experiment largely answered the research questions raised in the Introduction. We have seen clear results on whether and how boundary, focus, and newness affect F_0, and how they affect duration. These results are summarized briefly below, followed by an in-depth discussion.

(a) Does PFC apply across boundaries of different strengths? In particular, would a strong boundary block PFC?

Answer: Yes, PFC applies in all the boundary conditions, and it is not blocked by a strong boundary with a relatively long silent pause (longer than 200 ms). Moreover, F_0 of postfocus words is lowered in all the boundary conditions to roughly the same degree. Meanwhile, in-focus F_0 raising also applies in words either before or after a boundary. In short, boundary strength does not affect how focus is realized.

(b) Is boundary strength marked by both F_0 and duration? How are boundary strength and focus encoded simultaneously in intonation?

Answer: Duration is a more consistent cue of boundary strength, while the role of F_0 is limited and conditional. Phrase-final F_0 lowering and phrase-initial F_0 raising are applied only when the boundary is strong. In our case, it is between two clauses that are separated by a long silent pause. However, F_0 does not reflect different boundary degrees. Instead, the accumulative effect of preboundary lengthening and silent pause is sensitive to all the three boundary strengths. In addition, there is no consistent phrase-initial lengthening. The effect of boundary strength and focus in the phrase-final word occurs in parallel in that final F_0 lowering and in-focus F_0 raising both apply, and so do final lengthening and in-focus lengthening. However, interactions between boundary marking and focus still occur. A late focus in the second constituent smooths the connection between the two constituents by reducing the effect of final lowering of the first constituent.

(c) Does newness have any effect on PFC?

Answer: No, newness of postboundary words shows no systematic effect on F_0. As expected, pitch range of postfocus words is compressed regardless of whether they are given or new.

Altogether, the results are in favor of a functional view of focus assignment, i.e., speakers place focus on a part of an utterance they need to highlight based on the discourse context (Bolinger, 1972; Chafe, 1974; Halliday, 1967), whereas phrasing and newness of information do not change how focus is realized. Phrasing, on the other hand, is independent of focus, because its function is to signal the closeness of adjacent words, mostly via durational adjustment of preboundary words and optional silent pauses. Phrase-final F_0 lowering and phrase-initial F_0 raising are applied when a boundary is relatively strong, e.g. between two clauses and with a long silent pause. Note that two functions being independent of each other does not mean they do not interact. Rather, it only means that they convey different meanings and they have their own encoding mechanisms. In statistic terms, two factors have to be both independent before their interactions can be examined. Factors that are not independent of each other would in contrast be considered as being confounded.

4.1 Focus and Newness

The present results about focus realization are consistent with previous findings (Chen and Gussenhoven, 2008; Wang and Féry, 2012; Wang and Xu, 2011; Xu, 1999; Xu et al., 2012). The trizone pattern of focus is seen in all the boundary and newness conditions (Fig. 1, 2). That is, prefocus F_0 is largely intact, in-focus F_0 is raised and expanded, while postfocus F_0 is lowered and compressed. Importantly, PFC applies across a strong boundary, which is consistent with the findings about split sentences of Chinese (Wang and Féry, 2012). We can now generalize that a relatively long pause per se does not block PFC, based on two facts. First, the manifestation of focus is largely independent of other communicative factors, e.g., boundary marking in the present study. Second, the role of F_0 in signaling boundary strength is limited (see more detailed discussion in section 4.2).

Also, as mentioned in the Method section (2.1), when postboundary words are new, the type of focus can be described as corrective; whereas when the postboundary words are given, the focus type can be described as informational. The results here, however, show no difference between the given and new conditions (Fig. 4). This also means that the so-called contrastive focus and information focus do not differ in F_0. This finding is consistent with the finding of Kügler and Ganzel (2014) that there is no difference between corrective focus and counterpreposition focus in Mandarin, and the finding of Chen and Gussenhoven (2008) that exaggerated focus differs from regular focus in duration but not in F_0.

As expected, newness, as defined by whether a word has been mentioned in the previous context, did not show any effect on F_0 and duration. The newness of post-boundary words did not lead to any overall F_0 raising or give rise to any additional focus in the postfocus part, not even in the case when there is a long pause separating the focus and postfocus parts (Fig. 4). Katz and Selkirk (2011) have also found in English that words under contrastive focus show greater duration, intensity, and F_0 movement than elements that are discourse *new*. This indicates that a word is not necessarily focused just because it is first mentioned. Despite proposals that newness is one of the factors that determine the occurrence of focus (Brown, 1983; Chafe, 1976; Nooteboom and Kruyt, 1987; Nooteboom and Terken, 1982; Prince, 1981), the present results, consistent with those of Katz and Selkirk (2011), show that as long as focus is explicitly controlled, newness does not lead to additional variation in F_0 or duration. These results are also consistent with the finding for German that degree of newness is not reflected in F_0 (Baumann and Hadelich, 2003). As Ladd (1996) and Terken (1984) have suggested, there are factors other than newness that determine the occurrence of pitch accents. We can see that when focus and newness are defined separately as in the current study and in Chen et al. (2014), their phonetic manifestations as well as cognitive processing can be clearly distinguished.

4.2 Boundary Marking

In this study, we carefully controlled focus and newness of the postboundary words when examining the effect of boundary marking. It is surprising that, when all the other factors are largely controlled, the effect of boundary on F_0 is limited and sometimes even not sensitive enough to distinguish all the three boundaries. The boundary effect shows mostly in the preboundary word, in the form of phrase-final F_0 lowering at a strong boundary. Phrase-initial F_0 raising occurs only when a boundary is strong, and mostly only in terms of minimum F_0. In addition, the boundary effect on F_0 of the phrase-final word seems to be magnified by focus. When the preboundary or postboundary word is in focus, the preboundary word has lower F_0 in B3 conditions than in B1 and B2 conditions. The difference is 1.2 st on average. Such an effect disappears in neutral focus and Z focus (Fig. 5). A late focus or no focus seems to smooth the F_0 conjunction between two phrases. Interestingly, such an effect does not have any impact on pause duration (Fig. 6). Thus, the scope of the F_0 manifestation of focus can be as large as the whole sentence, while duration adjustment by focus is local, mostly on the focused word itself.

Another tonal effect worth mentioning is that lexical tone is more fully realized before a stronger boundary. The difference between B1 and B2 boundaries is noticeable just by looking at the tonal contours (Fig. 3). Tonal realization could also be used as a cue for boundary perception. We can notice that tones are more fully realized in focus as well. The difference between the full realizations of tone at a boundary and in focus is that focus causes large pitch raising and pitch range expansion of the whole tone and PFC, whereas boundary generates pitch lowering only in the final part of the tone.

Swerts (1997) reported that pitch reset is related to boundary strength in Dutch. Pitch reset, in that study, was measured in two steps. First, in any given phrase, the highest F_0 peak in an accented syllable at the vowel's amplitude maximum was taken as a measure of pitch range. Second, the distance in semitones was measured between the pitch range values before and after a given boundary of a particular strength. The correlation between boundary strength and pitch reset was, though significant, only 0.35. Although this measurement reduced some effect of focus-led pitch raising, it does not really reveal how a boundary affects F_0. From the current data, we can see that boundary effects are local, limited to only syllables adjacent to the boundary, and mostly in the preboundary syllable (Fig. 3). The large pitch reset

(>4.5 st) found by Swerts (1997) may very likely have come from the new topic effect (Umeda, 1982; Wang and Xu, 2011; Wang et al., 2011), given that the scope of the reset measurement covers a large temporal domain and the material used in that study.

We also notice that in previous studies, while talking about pitch reset at the prosodic boundary, hardly any direct comparison is made. But whenever it is possible to keep everything else constant, the boundary effect on F_0 is very small. For instance, Ladd (1988) used sentences in the form of "Allen is a stronger campaigner, and Ryan has more popular policies, but Warren has a lot more money." By changing "and" and "but" in the sentence, "A and B but C" can be directly compared with "A but B and C." His finding is that the topline of the B constituent starts with a higher pitch when it is after "but" than when it is after "and" (5-Hz difference on average for one speaker), and the difference only holds in the phrase-initial point, but not in any following pitch points or in preboundary ending pitch. We can see that although there is some difference in phrase-initial pitch between two boundary strengths, the difference is very small (Fig. 3 in Ladd, 1988), and even that difference may at least partially reflect a known effect of topic shift introduced by the word "but." Such a topic shift or new topic has been found to increase F_0 in sentence-initial words (Umeda, 1982; Wang and Xu, 2011; Wang et al., 2011). Thus, it is not easy to tease apart the two sources of phrase-initial F_0 raising, namely boundary marking and topic shift, while studying the boundary effect between sentences.

Some other similar experiments with controlled material led to the same finding that duration is a more consistent cue of boundary marking than F_0 (English: Allbritton et al., 1996; Katz et al., 1996; Lehiste, 1973; Mandarin: Xu and Wang, 2009). In the present study, we found that by combining the two durational cues, i.e., preboundary lengthening and silent pause, as proposed in Xu (2009), not only can the 3 boundaries be well distinguished, but also the degree of similarity between different boundaries is clearly marked. As we have seen, the difference between B1 and B2 boundaries is smaller than that between B1/B2 and B3 boundaries. In Figure 7, we can see that preboundary lengthening is much greater in B3 than in either B1 or B2, with the latter two having much smaller differences between them. The average preboundary word durations for B1, B2, and B3 are 384.8, 411.1, and 469.3 ms, respectively. Since B3 has a much longer silent pause, when silent pause is added, the duration correspondence to boundary strength becomes even more consistent.

For pause duration, focus does not show any effect. The newness of postboundary words shows effect on sentence set 1 but not on sentence set 2. Thus, the newness effect on pause duration is not consistent. Also pause duration differs significantly between the two sentence sets, which may be due to some unknown factors other than boundary strength, sentence length and syntactic structure.

With regard to phrase-initial duration, as analyzed in section 3.3.3, unlike in previous studies (Korean: Cho and Keating, 2001; English: Fougeron and Keating, 1997), we did not find consistent phrase-initial lengthening when the boundary was stronger. Here we notice that although Cho and Keating (2001), Fougeron and Keating (1997) and Keating et al. (2003) all found linguopalatal contact and duration of the phrase-initial consonant to be greater in the higher prosodic boundary condition, the methods they used were very different. Fougeron and Keating (1997) compared the phrase-initial syllable with all the other syllables in medial or final positions of the same prosodic category, e.g. intermediate phrase or intonational phrase.

In other words, it shows that a word at the phrase-initial position is longer than the same word at phrase-medial and -final positions. This is a *sequential comparison*. In contrast, Cho and Keating (2001) used an *in situ comparison* as in the current study. Different syntactic structures were used to control boundary strength while keeping words before the target word identical across different boundary strengths. They found that the initial consonant was longer when the boundary strength was greater. Keating et al. (2003) used a similar method to compare English, French, Korean and Taiwanese. Although they found that in all these languages speakers made some distinction on phrase-initial consonant between word-internal and phrase-initial conditions, there was no systematic cue for other prosodic levels, e.g., syllable-, word- or small-phrase-initial. In the current study, we found that not only phrase-initial duration was not used to distinguish boundary strength, but also phrase-initial F_0 showed no effect on boundary strength.

Putting all the results together, we can conclude that duration is the main acoustic cue for boundary marking. Final lengthening and optional pause duration are highly sensitive to boundary strength. The effect on F_0 by boundary strength is limited to final lowering at a strong boundary. Phrase-initial duration does not show any stable or systematic lengthening even at a strong boundary.

Although intensity has also been found to be important for marking boundary (Fougeron and Keating, 1997; Lehiste, 1973) and prominence (Kochanski et al., 2005), due to space limit, we did not discuss the role of intensity in this paper. Based on previous research, however, intensity change is largely consistent with F_0 change for focus (Xu, Chen and Wang, 2012) and topic (Alku et al., 2002; Grosz and Hirschberg, 1992) marking. Future studies can also look into possible roles of intensity for boundary marking.

5 Conclusions

By keeping as many known factors as possible under systematic control, and in particular by making in situ rather than sequential comparisons, we have found results that largely confirm the trizone focus realization that was established by Cooper et al. (1985) for English and Xu (1999) for Mandarin, among many others. For the first time, the current results have shown that PFC, the most consistent aspect of the trizone pattern, is not blocked by a strong boundary with a long silent pause (over 200 ms). Furthermore, in line with Lehiste (1973), Katz et al. (1996), Allbritton et al. (1996), and Xu and Wang (2009), the current results have shown that the combined duration of preboundary word and silent pause (Xu, 2009) provides the most reliable cue for boundary strength, while the role of F_0 in boundary marking is limited to phrase-final lowering and phrase-initial raising of minimum F_0 at a strong boundary – a boundary between two clauses in the current study. The reverse also appears to be true, i.e., focus does not seem to change the basic boundary marking strategy, with no impact on final lengthening or pause duration. The only noticeable effect of focus on boundary is that a late focus in a sentence seems to reduce phrase-final F_0 lowering, thus slightly weakening a strong boundary. Finally, like in Wang and Xu (2011), newness is again found to have little direct effect on F_0. Put together, the results provide clear evidence that focus and boundary marking are two separate communicative functions encoded largely in parallel.

6 Appendix: Reading Material

In the following the abbreviations are as below:
XF, focus on word X (the word right before the boundary)
YF, focus on word Y (the word right after the boundary)
ZF, focus on word Z (the last word)
NF, neutral focus
1sg, first person
2sg, second person
NOM, nominalizer
ASP, aspect marker
CL, classifier
COP, copula (shi4 in Mandarin)

The numbers of 1–4 in the Pinying layer stand for high, rising, low-dipping, and falling tone, respectively; 0 stands for neutral tone.

In the Chinese text, the underlined words are focused. In the Pinyin and English translation, words with all capital letters are focused.

Sentence Set 1-B3-Given

XF: 你是问，我买了什么，都送给毛奶奶了？ 我买了**柚栗**，都送给毛奶奶了。

ni3 shi4 wen4, wo3 mai3 le0 shen2me0, dou1 song4gei3 mao2nai3nai0 le0? wo3 mai3 le0 YOU4LI4, dou1 song4gei3 mao2nai3nai3 le0.

you COP ask, 1sg buy ASP what, all give-to Maonainai ASP? 1sg buy ASP NUT, all give-to Maonainai ASP.

Are you asking, what did I buy and give all of them to Maonainai? I bought NUTS and gave them all to Maonainai.

YF: 你是问，我买了柚栗，送没送给毛奶奶？ 我买了柚栗，**都**送给毛奶奶了。

ni3 shi4 wen4, wo3 mai3 le0 you4li4, song4 mei2 song4gei3 mao2nai3nai0? wo3 mai3 le0 you4li4, DOU1 song4gei3 mao2nai3nai3 le0.

you COP ask, 1sg buy ASP nut, give not give-to Maonainai? 1sg buy ASP nut, ALL give-to Maonainai ASP.

Are you asking, I bought nuts, did I give them to Maonainai or not? I bought nuts and gave them ALL to Maonainai.

ZF: 你是问，我买了柚栗，都送给谁了？ 我买了柚栗，都送给**毛奶奶**了。

ni3 shi4 wen4, wo3 mai3 le0 you4li4, dou1 song4gei3 shui2 le0? wo3 mai3 le0 you4li4, dou1 song4gei3 MAO2NAI3NAI3 le0.

you COP ask, 1sg buy ASP nut, all give-to whom? 1sg buy ASP nut, all give-to MAONAINAI ASP.

Are you asking, I bought nuts, whom did I give them to? I bought nuts and gave them all to MAONAINAI.

Sentence Set 1-B3-New

NF: 我要告诉你一件事。 我买了柚栗，都送给毛奶奶了。
wo3 yao4 gao4shu4 ni3 yi2 jian4 shi4. wo3 mai3 le0 you4li4, dou1 song4gei3 mao2nai3nai3 le0.
1sg want tell 2sg one CL matter. 1sg buy ASP nut, all give-to Maonainai ASP.
I need to tell you something. I bought nuts and gave them all to Maonainai.

XF: 不是甜橙。我买了**柚栗**，都送给毛奶奶了。
bu2 shi4 tian2chen2. wo3 mai3 le0 YOU4LI4, dou1 song4gei3 mao2nai3nai3 le0.
Not COP orange. 1sg buy ASP NUT, all give-to Maonainai ASP.
It was not orange. I bought NUTS and gave them all to Maonainai.

YF: 不是一半。我买了柚栗，**都**送给毛奶奶了。
bu2 shi4 yi2ban4. wo3 mai3 le0 you4li4, DOU1 song4gei3 mao2nai3nai3 le0.
not COP half. 1sg buy ASP nut, ALL give-to Maonainai ASP.
It was not half. I bought nuts and gave them ALL to Maonainai.

ZF: 不是李妈妈。我买了柚栗，都送给**毛奶奶**了。
bu4 shi4 li3ma1ma0. wo3 mai3 le0 you4li4, dou1 song4gei3 MAO2NAI3NAI3 le0.

not be Limama. 1sg buy ASP nut, all give-to MAONAINAI ASP.
It was not Limama. I bought nuts and gave them all to MAONAINAI.

Sentence Set 1-B2-Given

XF: 你是问，我买的什么都送给毛奶奶了？我买的**柚栗**都送给毛奶奶了。

ni3 shi4 wen4, wo3 mai3 de0 shen2me0, dou1 song4gei3 mao2nai3nai0 le0? wo3 mai3 de0 YOU4LI4, dou1 song4gei3 mao2nai3nai3 le0.

you COP ask, 1sg buy NOM what, all give-to Maonainai ASP? 1sg buy NOM NUT, all give-to Maonainai ASP.

Are you asking, what was that which I bought and gave all of them to Maonainai? The NUTS I bought were all given to Maonainai.

YF: 你是问，我买的柚栗送没送给毛奶奶？我买的柚栗**都**送给毛奶奶了。

ni3 shi4 wen4, wo3 mai3 de0 you4li4 song4 mei2 song4gei3 mao2nai3nai0? wo3 mai3 de0 you4li4 DOU1 song4gei3 mao2nai3nai3 le0.

you COP ask, 1sg buy NOM nut give not give-to Maonainai? 1sg buy NOM nut ALL give-to Maonainai ASP.

Are you asking, did I give the nuts that I bought to Maonainai or not? The NUTS I bought were all given to Maonainai.

ZF: 你是问，我买的柚栗都送给谁了？我买的柚栗都送给**毛奶奶**了。

ni3 shi4 wen4, wo3 mai3 de0 you4li4 dou1 song4gei3 shui2 le0? wo3 mai3 de0 you4li4 dou1 song4gei3 MAO2NAI3NAI3 le0.

you COP ask, 1sg buy NOM nut all give-to whom ASP? 1sg buy NOM nut all give-to MAONAINAI ASP.

Are you asking, whom did I give all the nuts that I bought? The nuts I bought were all given to MAONAINAI.

Sentence Set 1-B2-New

The context sentences in the new condition were the same as those in sentence set 1-B3-new condition. Below we only listed the sentences with Chinese characters and English translation to save some space.

NF: 我要告诉你一件事。我买的柚栗都送给毛奶奶了。

I need to tell you something. The nuts I bought were all given to Maonainai.

XF: 不是甜橙。 我买的**柚栗**都送给毛奶奶了。

It is not orange. The NUTS I bought were all given to Maonainai.

YF: 不是一半。 我买的柚栗**都**送给毛奶奶了。

It is not half. The nuts I bought were ALL given to Maonainai.

ZF: 不是李妈妈。 我买的柚栗都送给**毛奶奶**了。

It is not Limama. The nuts I bought were all given to MAONAINAI.

Sentence Set 1-B1-Given

XF: 你是问, 我买的什么兜送给毛奶奶了？我买的**柚栗**兜送给毛奶奶了。

ni3 shi4 wen4, wo3 mai3 de0 shen2me0 dou1 song4gei3 mao2nai3nai0 le0? wo3 mai3 de0 YOU4LI4 dou1 song4gei3 mao2nai3nai3 le0.

you COP ask, 1sg buy NOM what tote give-to Maonainai ASP? 1sg buy NOM NUT tote give-to Maonainai ASP.

Are you asking, what kind of tote that I bought was given to Maonainai? The "NUT" tote I bought was given to Maonainai.

YF: 你是问，我买的柚栗什么送给毛奶奶了？我买的柚栗**兜**送给毛奶奶了。

ni3 shi4 wen4, wo3 mai3 de0 you4li4 shen2me0 song4gei3 mao2nai3nai0 le0? wo3 mai3 de0 you4li4 DOU1 song4gei3 mao2nai3nai3 le0.

you COP ask, 1sg buy NOM nut what give-to Maonainai? 1sg buy NOM nut TOTE give-to Maonainai ASP.

Are you asking, what kind of thing with nut that I bought was given to Maonainai? The "nut" TOTE I bought was given to Maonainai.

ZF: 你是问，我买的柚栗兜送给谁了？我买的柚栗兜送给**毛奶奶**了。

DOI: 10.1159/000453082

ni3 shi4 wen4, wo3 mai3 de0 you4li4 dou1 song4gei3 shui2 le0? wo3 mai3 de0 you4li4 dou1 song4gei3 MAO2NAI3NAI3 le0.

you COP ask, 1sg buy NOM nut tote give-to whom ASP? 1sg buy NOM nut tote give-to MAONAINAI ASP.

Are you asking, whom did I give the “nut” tote that I bought? The “nut” tote I bought was given to MAONAINAI.

Sentence Set 1-B1-New

Again, context sentences in the new condition of B1 boundary are almost the same as those in sentence set 1-B3-new, except that in the YF condition, to make a contrast with “dou1 (tote)”, we used “bu4 (cloth)” in the context sentence.

NF: 我要告诉你一件事。 我买的柚栗兜送给毛奶奶了。

I need to tell you something. The “nut” tote I bought was given to Maonainai.

XF: 不是甜橙。 我买的**柚栗**兜送给毛奶奶了。

It is not orange. The NUT tote I bought was given to Maonainai.

YF: 不是布。 我买的柚栗**兜**送给毛奶奶了。

It is not cloth. The nut TOTE I bought was given to Maonainai.

ZF: 不是李妈妈。 我买的柚栗兜送给**毛奶奶**了。

It is not Limama. The “nut” tote I bought was given to MAONAINAI.

The base sentences of sentence set 2 in the three boundary conditions. The context sentences for focus conditions and newness were all constructed the same way as those for sentence set 1. For the interest of space, we do not list all the context sentences for sentence set 2:

B3: 我买了绿蛙，会送给毛奶奶。

wo3 mai3 le0 lv4wa1, hui4 song4gei3 mao2nai3nai0.

1sg buy ASP frog, will give-to Maonainai.

I bought a frog and will give it to Maonainai.

B2: 我买的绿蛙会送给毛奶奶。

wo3 mai3 de0 lv4wa1 hui4 song4gei3 mao2nai3nai0.

1sg buy NOM frog will give-to Maonainai.

The frog that I bought will be given to Maonainai.

B1: 我买的绿蛙烩送给毛奶奶了。

wo3 mai3 de0 lv4wa1 hui4 song4gei3 mao2nai3nai0 le0.

1sg buy NOM frog stew give to Maonainai ASP.

The frog stew that I bought was given to Maonainai.

7 Acknowledgment

We thank Wentao Gu and Caroline Féry for many constructive comments. We are very thankful to the two anonymous reviewers for their insightful suggestions and comments. We are thankful to Beijing Haitian Ruisheng Science Technology Ltd. for the recording program (AudiRec), especially to Yufeng Hao and Lin He. This work was supported by the Natural Science Foundation of China (60905062) and the Program for New Century Excellent Talents in University (NCET-12-0584) to the first author, and the National Science Foundation (NSF BCS-1355479) to the second author. The manuscript was accomplished while the first author was a visiting scholar at Harvard-Yenching Institute.

References

Alku P, Vintturi J, Vilkman E (2002): Measuring the effect of fundamental frequency raising as a strategy for increasing vocal intensity in soft, normal and loud phonation. Speech Commun 38:321–334.

Allbritton DW, McKoon G, Ratcliff R (1996): Reliability of prosodic cues for resolving syntactic ambiguity. J Exp Psychol Learn 22:714–735.

Baumann S, Becker J, Grice M, Mücke D (2007): Tonal and articulatory marking of focus in German. 16th Int Congr Phonet Sci, Saarbrücken, pp 1029–1032.

Baumann S, Hadelich K (2003): Accent type and givenness: an experiment with auditory and visual priming. 15th Int Congr Phonet Sci, Barcelona, pp 1811–1814.
Beckman ME (1996): The parsing of prosody. Lang Cognitive Proc 11:17–67.
Beckman ME, Pierrehumbert J (1986): Intonational structure in Japanese and English. Phonol Yearb 3:255–309.
Boersma P, Weenink D (2005): Praat: doing phonetics by computer. http://www.fon.hum.uva.nl/praat/.
Bolinger DL (1958): Intonation and grammar. Lang Lear 8:31–37.
Bolinger DL (1972): Accent is predictable (if you're a mind-reader). Language 48:633–644.
Botinis A, Fourakis M, Gawronska B (1999): Focus identification in English, Greek and Swedish. 14th Int Congr Phonet Sci, San Francisco, pp 1557–1560.
Brown G (1983): Prosodic structure and given/new distinction; in Cutler A, Ladd DR (eds): Prosody: Models and Measurements. Berlin, Springer, pp 67–77.
Bruce G (1983): Accentuation and timing in Swedish. Folia Linguist 17:221–238.
Chafe WL (1974): Language and consciousness. Language 50:111–133.
Chafe WL (1976): Givenness, contrastiveness, definiteness, subject, topics, and point of view; in Li CN (ed): Subject and Topic. New York, Academic Press, pp 25–53.
Chen L, Wang L, Yang Y (2014): Distinguish between focus and newness: an ERP study. J Neurolinguist 31:28–41.
Chen YY (2004): Focus and intonational phrase boundary in standard Chinese. IEEE Int Symp Chinese Spok Lang Processing (ISCSLP 2004), pp 41–44.
Chen Y-Y, Gussenhoven C (2008): Emphasis and tonal implementation in Standard Chinese. J Phon 36:724–746.
Chen Y-Y, Xu Y (2006): Production of weak elements in speech – evidence from F_0 patterns of neutral tone in Standard Chinese. Phonetica 63:47–75.
Cho T, Keating PA (2001): Articulatory and acoustic studies on domain-initial strengthening in Korean. J Phon 29:155–190.
Cho T, McQueen JM, Cox EA (2007): Prosodically driven phonetic detail in speech processing: the case of domain-initial strengthening in English. J Phon 35:210–243.
Cooper WE, Eady SJ, Mueller PR (1985): Acoustical aspects of contrastive stress in question-answer contexts. J Acoust Soc Am 77:2142–2156.
Cruttenden A (1997): Intonation. Cambridge, Cambridge University Press.
De Jong KJ (1995): The supraglottal articulation of prominence in English: linguistic stress as localized hyperarticulation. J Acoust Soc Am 97:491–504.
De Pijper JR, Sandeman AA (1994): On the perceptual strength of prosodic boundaries and its relation to suprasegmental cues. J Acoust Soc Am 96:2037–2047.
Dilley L, Shattuck-Hufnagel S, Ostendorf M (1996): Glottalization of word-initial vowels as a function of prosodic structure. J Phon 24:423–444.
Duez D (1982): Silent and non-silent pauses in three speech styles. Lang Speech 25:11–28.
Eady SJ, Cooper WE (1986): Speech intonation and focus location in matched statements and questions. J Acoust Soc Am 80:402–415.
Fant G, Kruckenberg A (1996): Quantal theory of speech timing. Speech, music and status report. Fonetik 96:41–44.
Féry C, Kügler F (2008): Pitch accent scaling on given, new and focused constituents in German. J Phon 36:680–703.
Féry C, Samek-Lodovici V (2006): Focus projection and prosodic prominence in nested foci. Language 82:131–150.
Fougeron C (2001): Articulatory properties of initial segments in several prosodic constituents in French. J Phon 29:109–135.
Fougeron C, Keating PA (1997): Articulatory strengthening at edges of prosodic domains. J Acoust Soc Am 101:3728–3740.
Fowler C, Housum J (1987): Talkers' signaling of "new" and "old" words in speech and listeners' perception and use of the distinction. J Mem Lang 26:489–504.
Grosz B, Hirschberg J (1992): Some intonational characteristics of discourse structure. 2nd Int Conf Spoken Lang Processing (ICSLP), Alberta, pp 429–432.
Gussenhoven C (2007): Types of focus in English; in Lee C, Gordon M, Büring D (eds): Topic and Focus: Cross-Linguistic Perspectives on Meaning and Intonation. Dordrecht, Springer, pp 83–100.
Halliday MAK (1967): Notes on transitivity and theme in English. Part II. J Linguist 3:199–244.
Hanssen J, Peters J, Gussenhoven C (2008): Prosodic effects of focus in Dutch declaratives. 4th Int Conf Speech Prosody, Campinas, pp 609–612.
He X, Hanssen J, van Heuven VJ, Gussenhoven C (2011): Phonetic implementation must be learnt: native versus Chinese realization of focus accent in Dutch. 17th Int Congr Phonet Sci, Hong Kong, pp 843–846.
Hirschberg J (1993): Pitch accent in context: predicting intonational prominence from text. Artif Intell 63:305–340.
Horne M, Strangert E, Heldner M (1995): Prosodic boundary strength in Swedish: final lengthening and silent interval duration. 13th Int Congr Phonet Sci, Stockholm, pp 170–173.
House J, Sityaev D (2003): Phonetic and phonological correlates of broad, narrow and contrastive focus in English. 15th Int Congr Phonet Sci, Barcelona, pp 1819–1822.
Hwang HK (2012): Asymmetries between production, perception and comprehension of focus types in Japanese; in Ma Q, Ding H, Hirst D (eds): Speech Prosody 2012, Shanghai, pp 641–644.

Ipek C (2011): Phonetic realization of focus with no on-focus pitch range expansion in Turkish. 17th Int Congr Phonet Sci, Hong Kong, pp 140–143.
Ipek C, Jun S-A (2013): Towards a model of intonational phonology of Turkish: neutral intonation. J Acoust Soc Am 133:3573.
Ishihara S (2002): Intonation of wh-questions in Japanese and its influence on syntax; in Otsu Y (ed): Linguistics and Phonetics 2002. Urayasu, Meikai University, pp 165–189.
Ishihara S (2011): Japanese focus prosody revisited: freeing focus from prosodic phrasing. Lingua 121:1870–1889.
Ito J, Mester A (2013): Prosodic subcategories in Japanese. Lingua 124:20–40.
Jeon H-S, Nolan F (2013): The role of pitch and timing cues in the perception of phrasal grouping in Seoul Korean. J Acoust Soc Am 133:3039–3049.
Jun S-A (1993): The Phonetics and Phonology of Korean Prosody; PhD thesis, Ohio State University, Columbus.
Katz J, Selkirk E (2011): Contrastive focus vs discourse-new: evidence from phonetic prominence in English. Language 87:771–816.
Katz WF, Beach CM, Jenouri K, Verma S (1996): Duration and fundamental frequency correlates of phrase boundaries in productions by children and adults. J Acoust Soc Am 99:3179–3191.
Keating PA, Cho T, Fougeron C, Hsu C-S (2003): Domain-initial articulatory strengthening in four languages. Lab Phonol 6:145–163.
Kochanski G, Grabe E, Coleman J, Rosner B (2005): Loudness predicts prominence: fundamental frequency lends little. J Acoust Soc Am 118:1038–1054.
Krifka M (2008): Basic notions of information structure. Acta Linguist Hung 55:243–276.
Krivokapic J, Byrd D (2012): Prosodic boundary strength: an articulatory and perceptual study. J Phon 40:430–442.
Kubozono H (1993): The Organization of Japanese Prosody. Tokyo, Kurosio Publishers.
Kubozono H (2007): Focus and intonation in Japanese: does focus trigger pitch reset. Interdisciplinary Studies on Information Structure (Working Papers of the SFB 632) 9:1–27.
Kügler F, Ganzel S (2014): On the elicitation of focus – prosodic differences as a function of sentence mode of the context? 4th Int Symp Tonal Aspects of Lang (TAL), Nijmegen, pp 71–74.
Ladd DR (1988): Declination “reset” and the hierarchical organization of utterances. J Acoust Soc Am 84:530–544.
Ladd DR (2008): Intonational Phonology. Cambridge, Cambridge University Press.
Ladd DR, Campbell N (1991): Theories of prosodic structure: evidence from syllable duration. 12th Int Congr Phonet Sci, Aix-en-Provence, pp 290–293.
Ladd R (1996): Intonational Phonology. Cambridge, Cambridge University Press.
Lee A, Xu Y (2012): Revisiting focus prosody in Japanese. Speech Prosody 2012, Shanghai, pp 274–277.
Lee Y-C, Xu Y (2010): Phonetic realization of contrastive focus in Korean. Speech Prosody, Chicago, pp 100031–100034.
Lehiste I (1973): Phonetic disambiguation of syntactic ambiguity. Glossa 7:107–122.
Lehiste I, Olive JP, Streeter LA (1976): Role of duration in disambiguating syntactically ambiguous sentences. J Acoust Soc Am 60:1199–1202.
Liberman M, Pierrehumbert J (1984): Intonational invariance under changes in pitch range and length; in Aronoff M, Oehrle RT (eds): Language Sound Structure. Cambridge, MIT Press, pp 157–233.
Lieberman P (1967): Intonation, Perception, and Language. MIT Research Monograph 38. Cambridge, MIT Press.
Liu F, Xu Y (2005): Parallel encoding of focus and interrogative meaning in Mandarin intonation. Phonetica 62:70–87.
Liu F, Xu Y, Prom-on S, Yu ACL (2013): Morpheme-like prosodic functions: evidence from acoustic analysis and computational modeling. J Speech Sci 3:85–140.
Nagahara H (1994): Phonological Phrasing in Japanese; PhD thesis, UCLA, Los Angeles.
Nespor M, Vogel I (1986): Prosodic Phonology. Dordrecht, Foris.
Nooteboom SG, Kruyt JG (1987): Accents, focus distribution, and the perceived distribution of given and new information: an experiment. J Acoust Soc Am 82:1512–1523.
Nooteboom SG, Terken J (1982): What makes speakers omit pitch accents? An experiment. Phonetica 39:317–336.
O’Connor JD, Arnold GF (1961): Intonation of Colloquial English. London, Longman.
Oller DK (1973): The effect of position in utterance on speech segment duration in English. J Acoust Soc Am 54:1235–1247.
Palmer HE (1922): English Intonation with Systematic Exercises. Cambridge, Heffer.
Pell MD (2001): Influence of emotion and focus on prosody in matched statements and questions. J Acoust Soc Am 109:1668–1680.
Pierrehumbert J (1980): The Phonology and Phonetics of English Intonation. Cambridge, MIT Press.
Pierrehumbert J, Hirschberg J (1990): The meaning of intonational contours in the interpretation of discourse; in Cohen PR, Morgan JL, Pollack ME (eds): Intentions in Communication. Cambridge, MIT Press, pp 271–311.
Pierrehumbert JB, Beckman ME (1988): Japanese Tone Structure. Cambridge, MIT Press.
Price P (1991): The use of prosody in syntactic disambiguation. J Acoust Soc Am 90:2956–2979.
Prince EF (1981): Toward a taxonomy of given-new information; in Cole P (ed): Radical Pragmatics. New York, Academic Press, pp 223–255.
Prince EF (1992): The ZPG letter: subjects, definiteness and information-status; in Mann WC, Thompson SA (eds): Discourse Description: Diverse Analyses of a Fund Raising Text. Amsterdam, Benjamins, pp 295–325.

Prom-on S, Xu Y, Thipakorn B (2009): Modeling tone and intonation in Mandarin and English as a process of target approximation. J Acoust Soc Am 125:405–424.
Rump HH, Collier R (1996): Focus conditions and the prominence of pitch-accented syllables. Lang Speech 39:1–17.
Sahkai H, Kalvik M-L, Mihkla M (2013): Prosody of contrastive focus in Estonian. 14th Interspeech, Lyon, pp 315–319.
Selkirk E (1986): On derived domains in sentence phonology. Phonol Yearb 3:371–405.
Selkirk E (2002): Contrastive focus vs presentational focus: prosodic evidence from right node raising in English; in Bel B, Marlin I (eds): Proceedings of the 1st International Conference on Speech Prosody, Aix-en-Provence. Aix-en-Provence, Laboratoire Parole et Langage, Université de Provence, pp 643–646.
Selkirk E (2005): Comments on intonational phrasing in English; in Frota, Sonia, Marina Vigário, and Maria João Freitas, (eds): *Prosodies: With special reference to Iberian languages*. Vol. 9. Walter de Gruyter, 2005. Berlin, Mouton de Gruyter, pp 11–58.
Selkirk E (2008): Contrastive focus, givenness and the unmarked status of "discourse-new". Acta Linguist Hung 55:331–346.
Selkirk E (2011): The syntax-phonology interface; in Goldsmith JA, Riggle J, Yu ACL (eds): The Handbook of Phonological Theory, ed 2. Oxford, Wiley-Blackwell, pp 435–484.
Selkirk E, Shinya T, Kawahara S (2004): Phonological and phonetic effects of minor phrase length on f0 in Japanese. 2nd Int Conf Speech Prosody, Nara, pp 345–348.
Shannon CE (1948): A mathematical theory of communication. Bell System Tech J 27:379–423.
Sugahara M (2005): Post-focus prosodic phrase boundaries in Tokyo Japanese: asymmetric behavior of an f0 cue and domain-final lengthening. Stud Linguist 59:144–173.
Swerts M (1994): Prosody as a marker of information flow in spoken discourse. Lang Speech 37:21–34.
Swerts M (1997): Prosodic features at discourse boundaries of different strength. J Acoust Soc Am 101:514–521.
Swerts M (1998): Filled pauses as markers of discourse structure. J Pragmat 30:485–496.
Swerts M, Geluykens R (1994): Prosody as a marker of information flow in spoken discourse. Lang Speech 37:21–43.
Swerts M, Krahmer E, Avesani C (2002): Prosodic marking of information status in Dutch and Italian: a comparative analysis. J Phon 30:629–654.
Swerts M, Strangert E, Heldner M (1996): F0 declination in spontaneous and read-aloud speech. Int Conf Spoken Lang Processing (ICSLP), Philadelphia, pp 1501–1504.
Terken J (1984): The distribution of pitch accents in instructions as a function of discourse structure. Lang Speech 27:269–289.
Terken J, Hirschberg J (1994): Deaccentuation of words representing "given" information: effects of persistence of grammatical function and surface position. Lang Speech 37:125–145.
Turk AE, Shattuck-Hufnagel S (2007): Multiple targets of phrase-final lengthening in American English words. J Phon 35:445–472.
Umeda N (1982): "F0 declination" is situation dependent. J Phon 10:279–290.
Vainio M, Mixdorff H, Järvikivi J, Werner S (2003): The production and perception of focus in Finnish. 15th Int Congr Phonet Sci, Barcelona, pp 1831–1834.
Wagner M (2005): Prosody and Recursion; PhD thesis, MIT, Cambridge.
Wang B, Féry C (2012): Prosodic encoding of topic and focus and its influence on speech perception in split sentences of Chinese Putonghua. Chin J Acoust 31:101–112.
Wang B, Xu Y (2011): Differential prosodic encoding of topic and focus in sentence-initial position in Mandarin Chinese. J Phon 39:595–611.
Wang B, Xu Y, Xu J (2011): Prosodic realization of discourse topic in Mandarin Chinese: comparing professional with non-professional speakers. 17th Int Congr Phonet Sci, Hong Kong, pp 17–21.
Wang L, Wang B, Xu Y (2012): Prosodic encoding and perception of focus in Tibetan (Anduo dialect). 6th Speech Prosody, Shanghai, pp 286–289.
Wells JC (2006): English Intonation PB and Audio CD: An Introduction. Cambridge, Cambridge University Press.
Wightman CW, Shattuck-Hufnagel S, Ostendorf M, Price PJ (1992): Segmental durations in the vicinity of prosodic phrase boundaries. J Acoust Soc Am 91:1707–1717.
Xu Y (1999): Effects of tone and focus on the formation and alignment of f0 contours. J Phon 27:55–105.
Xu Y (2005): Speech melody as articulatorily implemented communicative functions. Speech Commun 46:220–251.
Xu Y: ProsodyPro – A tool for large-scale systematic prosody analysis; Tools and Resources for the Analysis of Speech Prosody (TRASP 2013), pp 7–10. (Aix-en-Provence, France 2013).
Xu Y (2009): Timing and coordination in tone and intonation – an articulatory-functional perspective. Lingua 119:906–927.
Xu Y, Chen S-W, Wang B (2012): Prosodic focus with and without post-focus compression (PFC): a typological divide within the same language family? Linguist Rev 29:131–147.
Xu Y, Wang ML (2009): Organizing syllables into groups – evidence from F0 and duration patterns in Mandarin. J Phon 37:502–520.
Xu Y, Xu CX (2005): Phonetic realization of focus in English declarative intonation. J Phon 33:159–197.
Zhang XX, Wang B, Wu Q (2012): Prosodic realization of focus in statement and question of Tibetan. 13th Interspeech, Portland, pp 667–670.