CAMBRIDGE
UNIVERSITY PRESS

ORIGINAL ARTICLE

Focus interpretation in L1 and L2: The role of prosodic prominence and clefting

Mengzhu Yan[1,*] (iD), Paul Warren[2] (iD) and Sasha Calhoun[2] (iD)

[1]Huazhong University of Science and Technology, Wuhan, China and [2]Te Herenga Waka – Victoria University of Wellington, Wellington, New Zealand
*Corresponding author. Email: mengzhu_yan@hust.edu.cn

Abstract
Although the importance of prosody in processing information structure (IS) has been shown in many languages including English and Mandarin, the interacting effects of prosody with other linguistic systems, such as syntax, are relatively under-studied, especially in L2. This paper reports two question-answer appropriateness rating experiments, investigating intermediate-to-high proficiency Mandarin listeners' integration of prosodic prominence and clefting cues in the interpretation of focus in their L1 and L2 (English). It was found that prosodic prominence was more effective than clefting as a cue to focus in L1 Mandarin. However, clefting was as effective as prominence in L2, showing L1–L2 differences in integrating multiple cues. The findings are discussed in terms of Mennen's L2 intonation learning theory (2015) and Bates and MacWhinney's Competition Model (1989), which provide a framework for understanding difficulties in acquiring the use of particular cues in L2. The current study contributes to our limited knowledge of a crucial part of L2 learning: how L2 learners process IS.

In successful comprehension, listeners need to not only understand the words being said but also integrate what they hear into a shared model of the information established in the discourse. They need to understand how each new utterance relates to and updates that model, using the *information structure* (IS) of each utterance (e.g., see Calhoun, 2010; Krifka, 2008). IS describes the way in which information is packaged in a discourse to serve the purpose of communication (Chafe, 1976; Krifka, 2008). One key part of IS is *focus*, that is, the part of each utterance which updates an explicit or implicit "question-under-discussion" (QUD) in the immediate context (Roberts, 1996). For example, consider a context like (1) where a lost wallet has been found:

(1) Context: A young man lost his camera and wallet in the bus. The driver and a passenger found them.

CrossMark

(2) Who found the wallet?
(3) The **passenger** found the wallet (bold indicates prosodic prominence).
(4) * The passenger found the **wallet**.

The QUD is in (2). In response, the answer in (3), with prominence on *passenger* is appropriate, but (4), with prominence on *wallet* is not. The subject *passenger* is the focus, specifically a QUD-focus, as it answers or updates the QUD. Another function of focus is to indicate contextually relevant contrastive alternatives (Rooth, 1992), for example, the focus on *passenger* in (3) also implies there are other people who could have found the wallet, and these alternatives are relevant to the interpretation of the utterance. The two main functions of focus, reflected in the same word *passenger*, are best viewed as orthogonal to each other (Calhoun, 2010; Vallduví, 2016). Appropriate focus marking is important for coherence and efficient integration of information into the discourse model. Inappropriate focus marking has been shown to be a significant source of communication breakdown in L2 contexts (Jenkins, 2002).

Along with prosodic prominence, as shown in (3), certain syntactic constructions (e.g., the *it*-cleft) are also key markers of IS in many languages (Féry & Ishihara, 2016; Kügler & Calhoun, 2020). This poses challenges for L2 learners, as these cues are used differently, and with different weightings, in different languages. More generally, prosody is particularly challenging for L2 learners given the comparative difficulty in establishing relevant categories of prosodic structure and the pervasiveness of many-to-many mappings between form and function (e.g., see Mennen, 2015). There is little research on how L2 listeners use prosodic cues to process IS, although there is recent interest in this area (e.g., Chen & Lai, 2011; Foltz, 2020; Ge et al., 2020; Lee & Fraundorf, 2017). These studies have shown that learning speech cues to IS in L2 is challenging, especially as this is not usually explicitly taught in classrooms (Braun & Tagliapietra, 2011; Trouvain et al., 2007). Acquisition becomes even more complex when other cues are involved, such as syntactic cues. The interaction of multiple cues in IS processing remains unclear in L1 and L2.

Previous studies have found that L1 Mandarin and English listeners weight cues to focus differently, with prosodic cues being stronger than clefting in L1 Mandarin and the reverse in English (Calhoun et al., 2021; Chen et al., 2012; Greif & Skopeteas, 2021; Yan & Calhoun, 2020). To the best of our knowledge, however, no research has investigated how these interact in L2 processing. In the current study, we attempt to fill in the research gap by looking at how Mandarin learners of English use prosodic and clefting cues to focus when processing their L1 (Mandarin) and L2 (English) using the same question-answer appropriateness rating task. We draw on Bates and MacWhinney's (1989) Competition Model (CM) (see also Li & MacWhinney, 2013) to investigate the relative strength of prosodic and clefting cues (see section "The role of prosody and clefting in the processing of focus in L1 and L2"). Given differences between the languages in the hierarchy of cue strength in signaling focus, Mandarin learners of English need to acquire new cue strengths in English. We also draw on Mennen's (2015) L2 intonation learning theory which establishes dimensions of prosodic differences

between languages to predict difficulties in L2 intonation learning (see section "The role of prosody and clefting in the processing of focus in L1 and L2").

In section "Focus marking", we discuss how focus is marked by prosodic and clefting cues in Mandarin and English. In section "The role of prosody and clefting in the processing of focus in L1 and L2", we review relevant existing L1 and L2 studies. In section "The current study", we present two experiments examining the interacting effects of prosodic and clefting cues to focus by Mandarin listeners when processing their L1 and L2.

Focus marking

Prosodic focus marking

In Mandarin and English, the location of prosodic prominence is a primary cue to focus for listeners (e.g., Breen et al., 2010; Kügler & Calhoun, 2020; Lee et al., 2016). Prosodic prominence is realized with similar acoustic parameters in the two languages, typically a larger movement in F0, lengthened duration and increased intensity (Breen et al., 2010; Xu, 1999). However, the details differ across the two languages. In English, utterance-level prosodic prominence is realized through pitch accenting. The focused word typically carries the nuclear pitch accent, that is, it is the most prominent in the intonational phrase (Calhoun, 2010). In Figure 1, *parents* (top) and *children* (bottom) are foci, typically marked by H* or L+H* pitch accents in the ToBI system (Brugos et al., 2006). These pitch accents are realized with a (steep) rise in pitch on the stressed syllable of the focused word. Focus is also marked by reduced pitch range in the post-focal region as in Figure 1. In addition, the duration for these two words is longer than their unfocused counterparts, as indicated by the word boundaries. In Mandarin, localized pitch movement has a distinctive function at the word level, marking lexical tones. Utterance-level prosodic prominence is realized by expansion of the pitch range (Xu, 1999). Figure 2 shows expanded pitch range of the word 乘客 "passenger" when focused (top) compared to unfocused (bottom). Note that in both cases the lexical tone shape (rise-fall) is preserved: rising tone in the first syllable and falling tone in the second syllable.

It is important to note that prosodic prominence can mark both QUD-focus and contrastive-focus in English and Mandarin (e.g., Breen et al., 2010; Chen & Braun, 2006; Greif, 2010). However, it remains debated whether different types of prosodic prominence (emphatic prominence/L+H* or normal prominence/H*) are associated with distinct categorical meanings, for example, H* signals new information (noncontrastive, QUD-focus), while L+H* signals contrastive information (contrastive focus) (Pierrehumbert & Hirschberg, 1990). Another line of studies showed that there may not be a clear distinction between the two accent types with regard to signaling contrast (Hedberg & Sosa, 2008; Ladd & Schepman, 2003). In particular, Rooth (1992) made no distinction between different types of nuclear prominence—it is therefore assumed that any nuclear prominence can activate alternatives, that is, mark contrastive focus. Nevertheless, whether a focus has a contrastive reading is also affected by other factors like context: under a context that provides a contrastive interpretation (e.g., (1)), normal/noncontrastive prosodic prominence, such as an H* accent, may also render the focus contrastive.
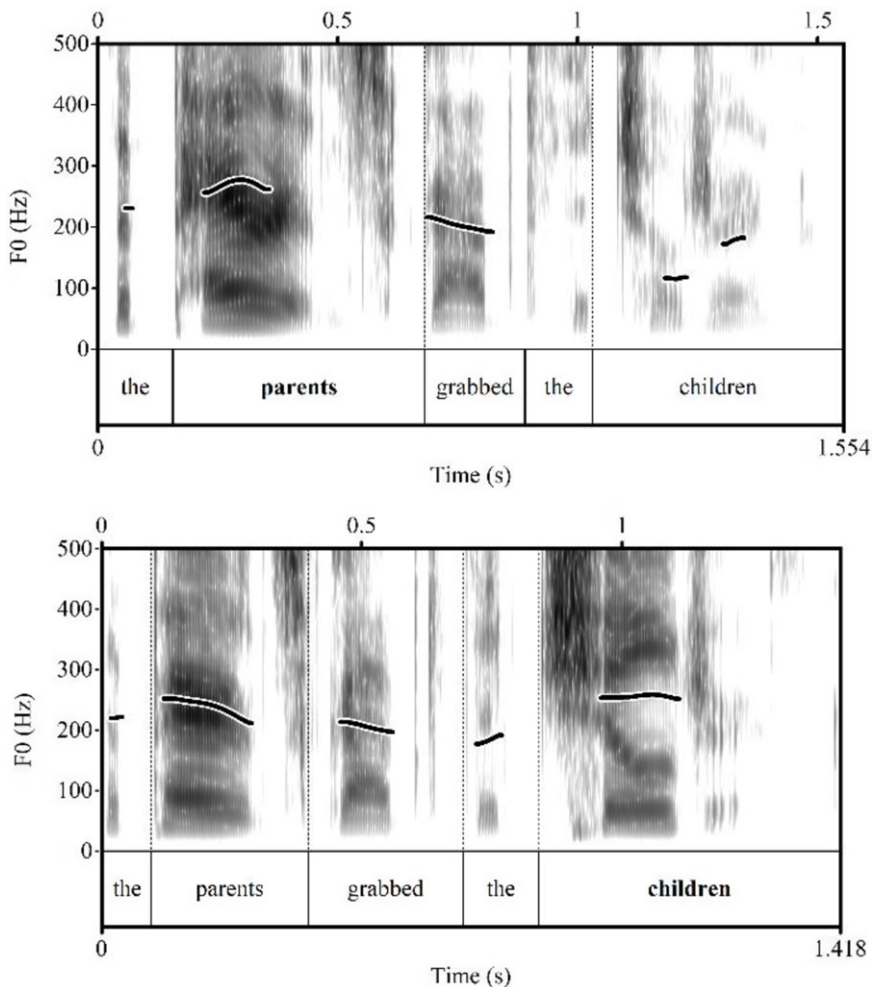
Figure 1. Prosodic prominence on the subject (top) and object (bottom) (taken from the English stimuli in the current study).

In both languages, sentence-final objects occupy the "default" focus position (Calhoun, 2010; Ladd, 2008; Xu, 2004). Subjects (and other non-sentence-final constituents) must carry phonetically strong prosodic marking to signal focus, with no prominence on the final object, while for final objects, "neutral" (non-emphatic) prominence is sufficient to signal focus. Further, objects have a default focus bias, that is, they may be processed as if they carry focus, even if they are not overtly focused marked at all (e.g., not accented, Carlson et al., 2009; Yan & Calhoun, 2020). Therefore, an asymmetry exists between subjects and objects in focus processing in both Mandarin and English (Chen et al., 2012; Yan & Calhoun, 2020).
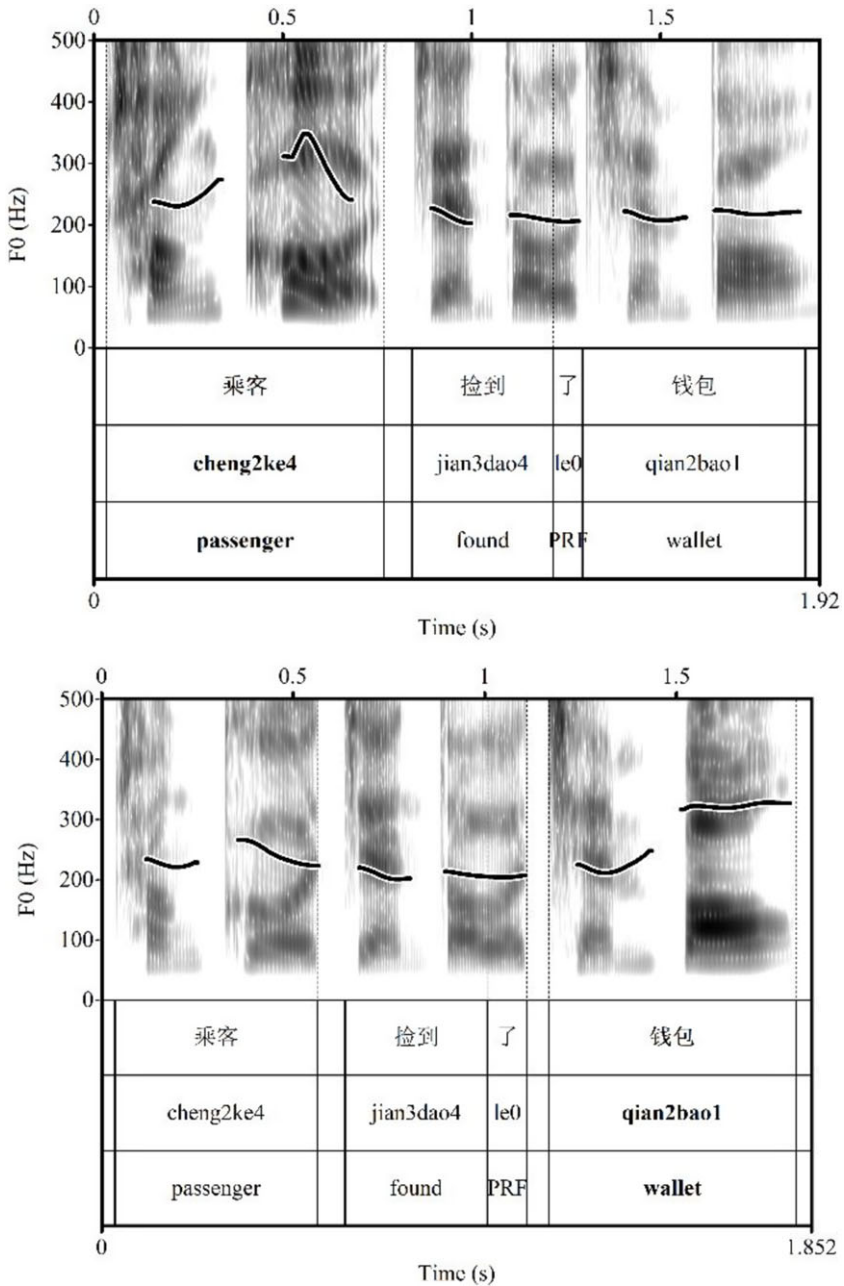
Figure 2. Prosodic prominence on the subject (top) and object (bottom); numbers in the second layer show tones (taken from the Mandarin stimuli in the current study).

Syntactic focus marking (clefting)

In English and Mandarin, focus can also be marked by certain syntactic constructions, such as clefts (e.g., *it was ... that* in English and *SHI....DE* in Mandarin) (Lambrecht, 2001; Paul & Whitman, 2008), as in the following:

(5) English
a. **Subject cleft**: It was [the **passenger**]$_F$ who found the wallet.
b. **Object cleft**: It was [the **wallet**]$_F$ that the passenger found.

(6) Mandarin
a. Subject cleft
是   [**乘客**]$_F$ 捡到   的   钱包
shi4   **cheng2ke4**   jian3dao4   de0   qian2bao1
SHI   **passenger**   found   DE   wallet
"It was [the **passenger**]$_F$ who found the wallet."

b. Object cleft
乘客   是   捡到   的   [**钱包**]$_F$
cheng2ke4   shi4   jian3dao4   de0   **qian2bao1**
passenger   SHI   found   DE   **wallet**
"It was [the **wallet**]$_F$ that the passenger found."

In both languages, the cleft head (e.g., *passenger* in (5a) and (6a)) is normally analyzed as being in focus. In English the head is always initial, immediately following *it was*, for example, *passenger* in (5a) and *wallet* in (5b). Mandarin has different positions for subject and object clefts. For subjects, *SHI* precedes the subject, as for 顾客 "passenger" in (6a), and *DE* precedes the object. *DE* in this position also indicates past tense (Chao, 1968; Hole, 2011; Paul & Whitman, 2008). For objects, *SHI* immediately precedes the verb, while *DE* precedes the object, as for 钱包 "wallet" in (6b) (Paul & Whitman, 2008). This pre-object 的 (DE) is largely restricted to Northern Mandarin speakers and may not be acceptable for Taiwan Mandarin and other southern varieties (Hole, 2011; Paul & Whitman, 2008; Teng, 1979), so the current study uses Northern Mandarin speakers.

SHI ... DE clefts in Mandarin and *it*-clefts in English are said to carry an existential presupposition that canonical sentences do not have (e.g., Hedberg, 2013; Lambrecht, 2001). For example, for (5a) and (6a), it is presupposed that someone (the passenger) found the wallet, and that there is an alternative set of people who could have. In order to be pragmatically appropriate, a context such as (1) (repeated as (7)) is usually needed, which introduces an entity that contrasts with the clefted element (Molnár, 2006). The contrasting entity can also be inferable from a set (e.g., *passenger*—people on a bus) (Calhoun et al., 2021). In the current study, such a context was included (see section "Materials and design").

(7) Context: A young man lost his camera and wallet in the bus. The driver and a passenger found them.

The clefted constituent is clearly therefore a contrastive focus. However, it is also usually a QUD-focus, in that it carries the asserted, or "question-answering"

part of the proposition (Lambrecht, 2001). This is consistent with the experimental findings for English and Mandarin (Arnhold, 2021; Calhoun et al., 2021; Chen et al., 2012) (see section "L1 studies"). It has also been claimed that another difference between the two types of focus marking, clefts and canonical word order sentences marked with prosodic prominence, is that clefts carry an exhaustive inference in both English and Mandarin (É Kiss, 1998; Hole, 2011; Krifka, 2008; Paul & Whitman, 2008). However, increasing experimental work in both languages is not consistent with this, unlike the exhaustive adverb *only* (e.g., Destruel et al., 2013, Liu & Yang, 2016; see overview in Onea, 2019). Reviewing these studies, Onea (2019) concludes that it is probably the case that "clefts imply that no other true answer to that issue is more informative than the canonical inference" (p. 415), which may or may not lead to the interpretation of exhaustivity depending on the context.

There are other types of cleft constructions in both languages, but we use *SHI . . . DE* clefts in Mandarin and *it*-clefts in English, since both are claimed to be canonical cleft constructions in the respective languages (Hole, 2012; Liu & Kempson, 2018) and have been studied for their focus properties. They are similar in function and have been compared in previous studies on focus processing (e.g., Yan & Calhoun, 2020).

In both languages, the cleft head normally carries the nuclear prominence, for example, 乘客 "passenger" in (5a) and (6a), so prosodic and clefting cues converge on the same word (which is both a QUD- and contrastive-focus). However, the nuclear prominence can fall on a different constituent, for example, the object 钱包 "*wallet*" in (8). In this case, prosodic and clefting cues mark QUD-focus on different constituents, that is, the two cues compete for focus assignment (both constituents are contrastive). Such "mismatch" constructions are attested in naturally occurring speech in both Mandarin and English (Delin & Oberlander, 1995; Hedberg, 2013; Hole, 2011; Lambrecht, 2001; Prince, 1978) and have recently been used in experimental studies of Mandarin, English, and Korean (Calhoun et al., 2021; Kember et al., 2021; Yan & Calhoun, 2019, 2020).

(8) Mismatch subject clefts with prosodic prominence on the object

English: It was [the passenger]$_F$ who found [the **wallet**]$_F$.
Mandarin: 是   [乘客]$_F$ 捡到   的   [钱包]$_F$
           shi4   cheng2ke4   jian3dao4   de0   **qian2bao1**
           SHI   passenger   found   DE   **wallet**
           'It was [the passenger]$_F$ who found [the **wallet**]$_F$.'

Despite their similarities, clefting cues to focus differ in Mandarin and English in the following ways. Firstly, the *SHI . . . DE* construction marks focus in Mandarin without word order alteration (other than the placement of *SHI*), while in English, clefts involve fronting (raising to a higher clause); this results in greater movement for the object, relative to its non-clefted position. Secondly, while subject clefts are more frequent than object clefts in both languages (Mandarin: Hedberg, 1999; English: Roland et al., 2007), this seems to be more marked in Mandarin, and object clefts may only mark object focus for Northern varieties of Mandarin (Paul & Whitman, 2008). These differences may affect their interpretation.

The role of prosody and clefting in the processing of focus in L1 and L2
L1 studies

In languages where prosodic prominence is a primary cue to focus, listeners use the position of prosodic prominence to identify the focus and distinguish focus types (e.g., Breen et al., 2010; Birch & Clifton, 1995; Clifton & Frazier, 2016; Lee et al., 2016; Welby, 2003). But there is much less research on what happens if there are multiple cues to focus in an utterance, particularly when these indicate focus on different words, for example, (8) above. In English, although prosodic prominence is a primary cue to focus in canonical order sentences, Calhoun et al. (2021) found listeners gave a higher weighting to the clefting cue (*it*-clefts) than prosodic prominence when the two cues signaled focus on different words. Calhoun et al. (2021) used a forced-choice task, in which participants had to choose between two possible questions for an utterance which was presented as an answer and which had prosodic and/or clefting cues on the subject and/or the object (the question and answer conditions were the same as those in the current study, see Table 1 in section "Materials and design"). They found that the preferences for choosing each question (subject question [SQ] or object question [OQ]) under different sentence conditions were as follows:

(9) SQ Preference: ScleftS≥canonS≥ScleftO
    OQ Preference: OcleftO≥canonO≥OcleftS

(canon = canonical order sentence; Scleft = subject cleft; Ocleft = object cleft; the capitalized S or O at the end of each label indicates whether prosodic prominence was on the subject or on the object, also see Table 1 in section "Materials and design" for detailed explanation of the labels. Note that in both cases ≥ indicates that the canonical sentences were not significantly different from either cleft, but these were significantly different from each other).

Calhoun et al. (2021) showed that prosodic prominence was a reliable cue to focus in canonical word order sentences. In cleft sentences where prosodic prominence and clefting cues aligned (ScleftS or OcleftO), the combined cue was key to listeners to interpret the focus position. However, when the two cues clashed (ScleftO or OcleftS), clefting was more important than prosodic prominence. For instance, "Who found the wallet?" was more likely to be chosen as a suitable question than "What did the passenger find?" for the answer "It was [the passenger]$_F$ that found [the **wallet**]$_F$." There were no differences in the effect of focus cues on subjects versus objects. Arnhold (2021) used a rating study (similar to the current study, see section "The current study") where listeners rated how suitable an answer sentence (canonical word order vs. subject clefts with/without prominence on the subject) was to the question (broad focus question vs. subject focus question). It was found that for the broad focus question the two clefts were rated less suitable than the canonical word order sentences, and the cleft with subject focus was rated less suitable than the one with broad focus. This suggests both prosodic and clefting cues played an important role in interpreting focus given a preceding broad focus question.

In Mandarin, studies suggest that prosodic prominence in Mandarin is more effective than clefting in the processing of focal information (Chen et al., 2012;

Yan & Calhoun, 2020). For example, Yan and Calhoun (2020) used a speeded "false alternative" rejection task, in which participants had to reject a "false alternative" question like (10) after hearing a sentence with the subject or object marked by prosodic prominence and/or clefting (as in (3)–(6), (8)) in Mandarin and English. They found that prosodic prominence facilitated the rejection of the false alternative (*driver* or *camera*) in both languages, with a stronger effect in Mandarin compared to English. Clefting played an inhibitory role in both languages. The results further showed that prosodic and clefting cues had a smaller effect on processing for objects than subjects, as focus was already cued by default position for objects (see section "Prosodic focus marking").

(10) SQ: Did the driver find the wallet?
      OQ: Did the passenger find the camera?

L2 studies

Before reviewing studies in L2, we present relevant L2 learning theories as a framework for evaluation of these findings. Few L2 learning theories explicitly address prosody. We concentrate on the L2 intonation learning theory (Mennen, 2015) and the CM (Bates & MacWhinney, 1989) as the most useful for the current study.

L2 learning theories

Mennen's L2 intonation learning theory (LILt) (2015) explicitly deals with intonation acquisition. Although the focus of the theory is on production, it has also been tested in perception studies (see section "L2 research on focus processing"). The theory predicts that L2 difficulties in learning intonation will depend on the types of differences between L1 and L2, due to L1 transfer. Four dimensions of differences (modified from Ladd, 1996) are presented in the model and discussed in relation to focus marking in English and Mandarin below.

The first is the "systemic dimension," involving the inventories of phonological elements and their distributions in the two languages. English and Mandarin seem to differ on this dimension in that focus is marked by pitch accenting in English and pitch range expansion in Mandarin. However, previous perception studies (see above and Ip & Cutler, 2020; Yan et al., 2022) have shown that these are functionally equivalent in processing the focus of an utterance. Therefore, we treat these as the same phonological element: nuclear (or utterance-level) prosodic prominence, in both languages. This places the difference on the second "realizational dimension" (the phonetic implementation of phonological elements: pitch accenting vs. pitch range expansion). We show below that this difference does not seem to matter in perception, at least for focus; however, it may be important in production. Therefore, it is predicted that the acquisition of prosodic prominence in perception may not be difficult for Mandarin learners of English due to the similar functionality of prosodic prominence in learners' L1 and L2. The third "semantic dimension" (relating to the functions of the categorical elements or tunes) is particularly relevant for the current study. In our case, this concerns how prosodic prominence interacts

with clefting as cues to focus in Mandarin and English. Although the model does not touch on how prosody interacts with other parts of the linguistic system, we predict difficulties for L2 learners if this interaction is functionally different to their L1 system. The languages seem to differ in the relative importance of prosodic and syntactic cues to focus. Based on LILt, we predict that the use of prosodic cues in English should not be problematic for Mandarin L2 listeners, but clefting may not provide as effective a cue to focus for Mandarin listeners in English. The fourth "frequency dimension" (the frequency of use of phonological elements) is not as easy to assess, because the relative frequencies of use of prosodic prominence and clefting as focus markers are not known (e.g., from corpora).

The CM (Bates & MacWhinney, 1989), as applied to second language acquisition (Li & MacWhinney, 2013), has been tested in much crosslinguistic research with both L1 and L2 (e.g., Chen et al., 2012; Liu et al., 1992). Although it is not specifically directed toward L2 learning of prosody, it can be used to make relevant predictions for the current study as the model deals with cue weights in learners' L1 and L2, which is the core of the current study. The weighting of various cues to signal meanings may differ between languages. The model suggests that while native speakers of a given language use the cue strength associated with that language, for learners of that language, the acquisition of cue-strength patterns is part of their learning and may exhibit different patterns depending on the L1–L2 relationship. In English and Mandarin, prosodic prominence and clefting, the two types of focus cues used in the current study, have different cue strengths. These cues converge or compete in the interpretation of focus. For example, in "It was the **passenger** who found the wallet," both the prosodic and clefting cue converge on *passenger* as the focus. The cue strength is combined. By contrast, the two cues compete in "mismatch" clefts such as "It was the passenger who found the **wallet**." The two cues indicate focus on different words, and this divergence creates a competition: prominence wins in Mandarin and clefting may win in English (e.g., Calhoun et al., 2021; Chen et al., 2012). In light of this variation in relative cue strength, when Mandarin learners interpret focus in their L2 English, they need to acquire new cue-strengths and L1 transfer is likely.

L2 research on focus processing
The use of prosodic cues to process focus in L2 speech is challenging, even for proficient learners (e.g., Braun & Tagliapietra, 2011; Ge et al., 2020; Lee & Fraundorf, 2017; van Maastricht et al., 2016). For example, using a cross-modal lexical priming task, Braun and Tagliapietra (2011) looked at the effect of contrastive and noncontrastive prosodic prominence in priming contrastive alternatives in Dutch (L2) with advanced L1 German listeners. Contrastive prosodic prominence enhances the encoding of alternatives to the focused word but is marked by different intonation contours in Dutch and German. It was found that learners with high proficiency inferred alternatives after both intonation contours, suggesting acquisition of the L2 contour but also L1 transfer.

Recently, Ge et al. (2020) investigated the mapping between focus and prosody in English sentences with *only* by advanced Cantonese and Dutch learners. Cantonese marks focus primarily with focus particles (Chao, 1947), whereas English and Dutch

use prosodic prominence. Ge et al. (2020) manipulated question (e.g., "I wonder what the fox is licking"[1]) and answer types (e.g., "The fox is only licking the **honey/licking** the honey") and found that Cantonese learners of English gave equal "make sense" ratings to appropriate prosody and to inappropriate prosody, while both English native speakers and Dutch learners of English gave higher "makes sense" ratings to answers with appropriate than inappropriate prosody. This can be explained by LILt as differences in the semantic dimension. Difficulties occurred when the use of prosodic cues to focus is different in L1 and L2, as was the case for Cantonese learners but not for Dutch learners. Ortega-Llebaria and Colantoni (2014) investigated the mapping between focus, prosody, and syntax in English with advanced Mandarin and Spanish learners, finding that in the perception and production of English focus using prosodic prominence, Mandarin learners behaved similarly to English native speakers, but Spanish learners performed less accurately. This again shows L2 performance largely depended on the semantic function of the L2 cues in their L1, consistent with LILt.

Although the mapping between prosody and focus has been investigated in L2 studies, the mapping between clefting and focus has rarely been tested (but see Reichle & Birdsong, 2014 for an investigation in L2 proficiency effects using event-related potentials with L1 and L2 French speakers), let alone the interaction between prosodic and clefting cues. Due to the differences in the relative cue weightings in the two languages, it is valuable to examine how clefting cues are used in focus identification in English by Mandarin learners. The previous studies reviewed above have used moderate to high proficiency learners, yet find processing differences compared with L1 speakers, suggesting this is a challenging aspect of second language acquisition even for reasonably advanced learners.

The current study

The current study investigates which cue(s) (prosodic, clefting, or both) Mandarin learners of English use in their interpretation of focus in an untimed rating task in their L2, compared to L1. In the experiments, participants judged how appropriate an answer to a question sounded. The intended focus in the answer was marked by prosodic and/or clefting cues. Given the contexts in the experiment, the focus was a narrow QUD-focus, that is, the focused word answered the preceding question (see section "Focus marking"). Both subject and object focus conditions were used, as previous studies have found differences in the processing of subjects and objects in reaction time and accuracy tasks (though not rating tasks, see section "L1 studies") and there are asymmetries between subject and object clefts (see section "Syntactic focus marking (clefting)"). We did not include a broad focus or "neutral" realization given the complexity of the design.

There were two groups of participants: near-monolingual Mandarin-speaking participants in their L1 (L1 experiment) and Mandarin learners of English (mid-to-high proficiency) in their L2 English (L2 experiment). Our study therefore focuses on intermediate-to-high level English learners. While proficiency is likely to affect performance on this task, we chose to look at these learners, as integration of multiple cues to focus is complex and likely to be challenging even for high

proficiency learners. This may be too difficult for lower proficiency learners. Previous studies in this area have involved learners of similar proficiency levels to ours (see section "L2 research on focus processing").

The L2 English materials were slightly simplified from those used in Calhoun et al. (2021) with L1 English speakers, using the same conditions. Some words in the materials in Calhoun et al. (2021) were replaced so that all words were among the 2000 most frequent words in BNC/COCA (Cobb, 2000), making sure that L2 learners knew them. This allowed us to compare our findings for L2 English speakers to those for L1 English speakers in Calhoun et al. (2021).

Research questions and predictions

Research questions

RQ1. Does prosodic prominence on the QUD focus result in higher appropriateness ratings than prosodic prominence in another location, and does this differ in L1 and L2?

RQ2. Relative to canonical word order, does consistent clefting on the QUD focus result in higher appropriateness ratings and inconsistent clefting in lower appropriateness ratings, and does this differ in L1 and L2?

RQ3. What are the relative weightings of prosodic and clefting cues in L1 and in L2?

Predictions

For RQ1, our prediction in both L1 and L2 is that prosodic prominence on the QUD focus will result in higher ratings than prosodic prominence in another location because of the key role of prominence in marking focus in both languages, as prosodic prominence cues to focus are used in a functionally equivalent way in learners' L1 and L2 based on the first two dimensions in LILt (see section "L2 learning theories"). For RQ2, in both L1 and L2, based on the CM (see section "L2 learning theories"), we predict that compared to no clefting, consistent clefting will strengthen focus marking and result in higher ratings, while inconsistent clefting cues will result in lower ratings. For RQ3, we expect different cue-weighting in L1 and L2: when prosodic and clefting cues compete, prosodic cues will win over clefting in Mandarin, but clefting may win over prosodic cues in English based on previous studies (see section "L1 studies"). Therefore, in L2, based on the "semantic dimension" in LILt (see section "L2 learning theories"), it is expected that effective use of clefting cues will be difficult for Mandarin learners, due to L1 transfer. Also, due to the possible difference in relative cue strength in learners' L1 and L2, the CM predicts that acquiring new cue strengths may pose challenges for L2 learners.

Method

Participants

Thirty-six native near-monolingual Mandarin speakers (32 females and 4 males; mean age = 24.6 [18–40], *SD* = 5.4) were recruited for the L1 experiment. We recruited L1 speakers who were similar to the L2 speakers specified below

(in terms of demographics, education level, etc.). Given that some English education is compulsory in China, it is highly unlikely that this group would be fully monolingual. All of our L1 participants had completed compulsory education in China and had therefore learnt English for at least 6 years. Their average self-reported English proficiency level was 2.7 ($SD = 0.6$) on a scale of 1–6. They had not lived outside China for more than six consecutive months.

Another 36 native Mandarin speakers (35 females and one male; mean age = 21 [19–23], $SD = 0.8$) were recruited for the L2 experiment. They were third or fourth year English major students recruited from a university in China. Their average self-reported English proficiency level was 3.85/6 ($SD = 0.6$), which was significantly higher than that of the participants for the L1 experiment ($t = 8.5$, $p < .001$). This was slightly lower than 4.17/6 for Cantonese learners and 4.55/6 for Dutch learners, who were categorized as advanced learners in Ge et al. (2020), and slightly higher than 3.6/6 in Lee and Fraundorf (2017) and 3.7/6 in Lee and Fraundorf (2019) for Korean learners who were categorized as moderate to high proficiency learners. Like the L1 participants, none of the L2 participants had lived outside China for more than six consecutive months. All participants in both experiments reported themselves to have grown up in northern or central China (e.g., Henan, Hebei province) and could not speak any language other than their native language (Mandarin and/or a dialect) and English. They received supermarket vouchers in recognition of their participation. None of them reported hearing or reading difficulties.

A cloze test, developed for another L2 focus processing experiment by Lee and Fraundorf (2019), was used to assess proficiency of the L2 group and was included as the proficiency factor in the data analysis (see section "Results"). The average cloze score was 31/40 ($SD = 3$) (as compared to 33.6 [$SD = 2.91$] in Lee & Fraundorf, 2019). The average age of starting to learn English was 10.7 ($SD = 2$). The means of self-reported English proficiency on a scale of 1–6 are listening: 3.4 ($SD = 0.9$), speaking: 3.5 ($SD = 1$), reading: 4.5 ($SD = 0.8$), and writing: 4 ($SD = 0.7$).

Materials and design

Forty-eight sets of Mandarin experimental materials for the L1 experiment and 60 sets of English materials for the L2 experiment were constructed. The Mandarin stimuli were adapted from Yan and Calhoun (2020) which used cross-modal lexical decision tasks to test the role of focus in speech processing in the two languages, where a lot of factors (e.g., frequency, the numbers of strokes and radicals of Mandarin characters) were controlled for. Each set of materials in the two languages includes a *Context*, *Focus probe question*, *Critical sentence (Answer)*, and *Rating question*, as shown in Table 1 (see full list in the On-line Supplementary Materials available at https://osf.io/54hgr/). The context was designed to be short and plausible, while introducing either two sets of explicitly mentioned alternatives (e.g., *passenger* and *wallet* in Table 1) or a set which introduced inferable alternatives (e.g., *man* and *woman* can be implied by *couple*). The alternatives in the context made the cleft sentences pragmatically plausible (see section "Syntactic focus

Table 1. Examples of test materials

Context		司机和一个乘客找到了那位年轻男子丢失的相机和钱包。 A young man lost his camera and wallet on the bus. The driver and a passenger found them.
Focus probe question	SQ	谁捡到了钱包？
		Who found the wallet?
	OQ	乘客捡到了什么？
		What did the passenger find?
Critical sentence (Answer)	canonS	[乘客]$_F$捡到了钱包 ([**S**]$_F$ V O)
		[The **passenger**]$_F$ found the wallet.
	canonO	乘客捡到了[**钱包**]$_F$ (S V [**O**]$_F$)
		The passenger found [the **wallet**]$_F$.
	ScleftS	是[**乘客**]$_F$捡到的钱包 (*SHI* [**S**]$_F$ V *DE* O)
		It was [the **passenger**]$_F$ who found the wallet.
	ScleftO	是[乘客]$_F$捡到的[**钱包**]$_F$ (*SHI* [S]$_F$ V *DE* [**O**]$_F$)
		It was [the **passenger**]$_F$ who found [the **wallet**]$_F$.
	OcleftS	[**乘客**]$_F$是捡到的[钱包]$_F$ ([**S**]$_F$ *SHI* V *DE* [O]$_F$)
		It was [the wallet]$_F$ that [the **passenger**]$_F$ found.
	OcleftO	乘客是捡到的[**钱包**]$_F$ (S *SHI* V *DE* [**O**]$_F$)
		It was [the **wallet**]$_F$ that the passenger found.
Rating question		答案听起来有多恰当？ How appropriate does the answer (to the question) sound?

marking (clefting)”). All words in the English critical sentences were among the 2000 most frequent words in BNC/COCA (Cobb, 2000).

As Table 1 shows, there were two types of question: a SQ and an OQ. There were six sentence types for the answer, varying the position of prosodic prominence and the presence and type of clefting: canonS and canonO (canonical word order with contrastive prosodic prominence on the subject or the object noun), ScleftS and ScleftO (subject cleft with contrastive prosodic prominence on the subject or the object noun), and OcleftS and OcleftO (object cleft with contrastive prosodic prominence on the subject or the object noun). This resulted in 12 experimental conditions (2 question types × 6 sentence types).

Each critical sentence appeared in all 12 conditions, giving a total of 576 experimental stimuli for the L1 experiment and 720 for the L2 experiment. Twelve lists of stimuli were constructed in a Latin square design so that each sentence was in a different condition in each list. Each participant saw one list.

A further 24 filler trials for the L1 experiment and 20 fillers for the L2 experiment were constructed. There were fewer fillers in the L2 experiment so as not to make it too long for participants. The focus probe questions in the fillers asked about different parts of the answer such as adverbial (e.g., “When did the scholar visit this city?”). The answers in the filler trials had different sentence structures from the test
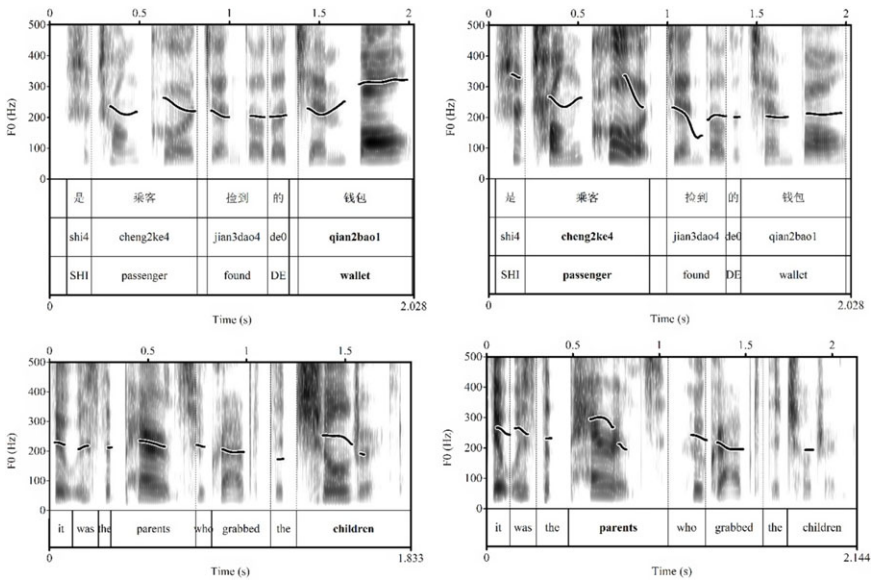
Figure 3. Example pitch tracks for the stimuli in the subject cleft conditions, with stress on the object (ScleftO, left) then subject (ScleftS, right). Mandarin is at the top, English bottom.

items, such as subject-verb and subject-adverbial-verb-object, with varied locations of prosodic prominence.

Recording and acoustic analysis

For each language, the focus probe questions (see Table 1) were recorded by a male native speaker and the critical sentences (see Table 1) by a female native speaker. The English speakers had North American varieties, which would be familiar to most Chinese learners of English. All speakers were given instructions about how to read the sentences. Recordings were made in a soundproof room via a head-mounted microphone directly to a computer hard drive using Praat (Boersma & Weenink, 2018) (Mandarin: sampling rate: 32000 Hz; bit rate: 512 kbps, bit depth: 16; English: sampling rate: 44100 Hz; bit rate: 705 kbps; bit depth: 16) (see Figure 1 for English examples of canonS and canonO, and Figure 2 for Mandarin examples of canonS and canonO, Figure 3 for examples of ScleftO and ScleftS in Mandarin and English, and Figure 4 for examples of OcleftO and OcleftS in Mandarin and English). The intensity of all sentences was re-scaled to 50 dB. Contrastive prosodic prominence (or emphatic stress) was used in the answer sentences in the experiments, as we wanted stimuli to unambiguously mark prosodic focus on either the subject or the object, and a "neutral" prominence (or normal stress) could leave the focus ambiguous.

To confirm that the critical sentences did indeed differ according to the intended position of the prosodic prominence, the Mandarin sentences were automatically segmented (with manual corrections) at the word level using the Montreal Forced Aligner (McAuliffe et al., 2017) and the English critical sentences using
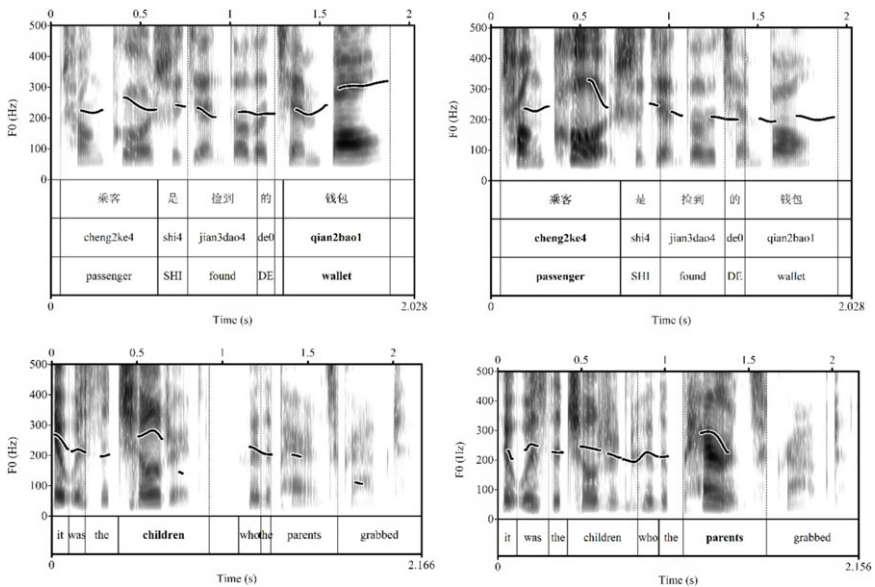
Figure 4. Example pitch tracks for the stimuli in the object cleft conditions, with stress on the object (OcleftO, left) then subject (OcleftS, right). Mandarin is at the top, English bottom.

Munich automatic segmentation (MAUS) (Kisler et al., 2017). Then acoustic measurements were obtained for the subject and object nouns. The measures were duration, mean F0, and mean intensity (see On-line Supplementary Material). As focus is marked through pitch range expansion in Mandarin, F0 range (max F0 – min F0) was also calculated for Mandarin. The differences between each word (subject and object) in each condition, and between the same word in each prominence condition (e.g., subject in canonS vs. canonO), confirmed the materials have the intended prominence locations in both languages. Statistical tests confirming these findings are available in the On-line Supplementary Materials.

Procedure

The L1 and L2 experiments were administered using OpenSesame v. 3.1 (Mathôt et al., 2012) in a quiet computer room. Participants listened to the stimuli over closed-ear headphones. Participants received written instructions on the computer screen. As Figure 5 shows, participants first saw a context and were instructed to press any key to proceed when they had read it, with no time limit. They then heard a dialogue consisting of a question in the male voice followed after 500 ms by an answer in the female voice. The screen was blank with a black background while the audio played. After that, they were prompted by a message on the screen to judge the appropriateness of the answer to the question, using a scale from 1 to 7. Only the endpoints (1 一点儿也不恰当 [not appropriate at all] and 7 极其恰当 [extremely appropriate]) were labeled. They indicated their judgment using the appropriate key on the keyboard, within 6 s. After either a key press or 6 s the experiment moved to
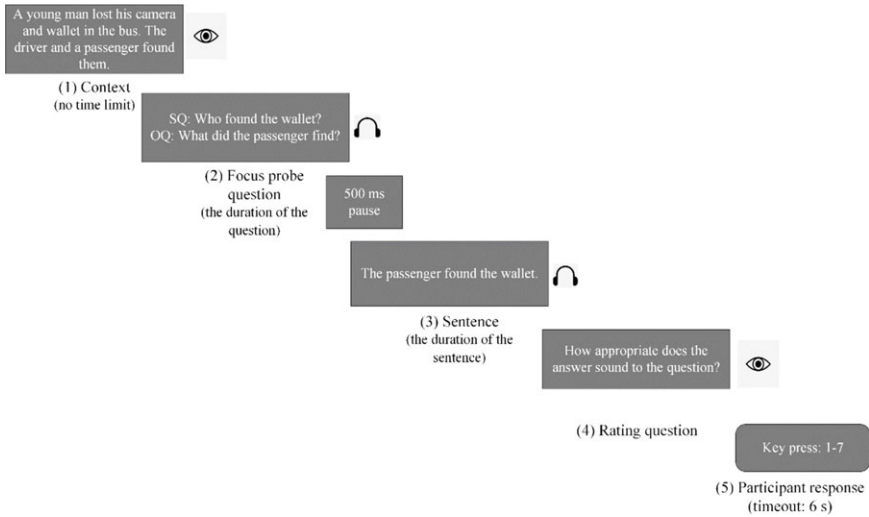
Figure 5. Procedure of the experiments.

the next trial. Trial order was randomized for each participant. Participants were told that they could take a break if they wanted whenever they saw the screen showing the context.

Six practice trials were played in a fixed order before the main experiment. These followed the same format as the main experiment. Demographic information was collected using a paper form at the end. The cloze test was then administered to measure L2 participants' English proficiency for the L2 experiment participants. The entire Mandarin experiment lasted approximately 20 min and the English experiment 30 min.

Data cleaning and analysis

The rating responses (1–7) were recorded and analyzed. For the L1 experiment, 50 of 1728 (48 from each of the 36 participants) test trials were unanswered (2.9%). The remaining 1678 test trials were used in the following analysis. For the L2 experiment, all 2160 test trials were used in the analysis.

As the dependent variable (ratings) was ordinal, cumulative link mixed models were fitted using the *ordinal* package (Christensen, 2015) in R (R Core Team, 2018). Because the L1 and L2 experiments involved different languages and groups of participants, it would not be appropriate to include both sets of data in one model. Therefore, we carried out two separate analyses. For the L1 experiment, the initial model included key experimental predictors: prosodic prominence position (subject, object), syntactic type (canonical, subject cleft, object cleft), and question type (subject, object) and their interactions. For the L2 experiment, we further included the values from the cloze test. This measure was a continuous variable in the four-way interaction (syntax*prosodic prominence*question type*proficiency[2]), in lower-level interactions, and as a simple effect. Random effects included

Table 2. ANOVA table of the final model for ratings for L1

	Chisq	df	p
Syntax	30	2	<.001
ProsodicProminence	5	1	.029
QuestionType	7	1	.01
Syntax:ProsodicProminence	20	2	<.001
Syntax:QuestionType	60	2	<.001
ProsodicProminence:QuestionType	1128	1	<.001
Syntax:ProsodicProminence:QuestionType	30	2	<.001

Note. Model: Syntax*ProsodicProminence*QuestionType + (1|Item) + (1 + Syntax|Participant)

intercepts for participants and items, and by-participant and by-item random slopes for the interactions between the key experimental factors as well as for their simple effects.

In each case, the initial model was simplified by removing the random slopes that had the lowest variance scores until the model converged. The R function *anova* was used to eliminate non-significant random slopes by comparison of models with and without one slope, and the *Anova.clmm* function in the *RVAideMemoire* package (Hervé, 2015) to eliminate non-significant fixed effects.

Results

L1 interpretation of focus

Table 2 shows the significance of variables in the final model for L1 (the analysis code, data, and the results of the fixed effects are summarized in the On-line Supplementary Materials). Predicted ratings were calculated from the model and plotted using the *ggplot2* package (Wickham et al., 2016) in Figure 6.

A compact letter display (CLD) was produced using the CLD function in the *multcompView* package (Graves et al., 2019) and added to Figure 6. This function produces all pairwise comparisons of least-squares means with *Tukey* adjustment (see the estimates, standard errors, z, and p values in the On-line Supplementary Material). The difference between two conditions is not significant ($p \geq .05$) when they share at least one letter. For example, the first two conditions in the figure have different letter labels (*e* and *d*) showing that they were significantly different from each other, while the first two conditions in black share the label *b* and were therefore not significantly different.

Figure 6 shows the following preferences of each answer type for SQ and OQ in the L1 experiment:

(11) SQ: canonS = ScleftS > OcleftS > canonO = ScleftO = OcleftO
OQ: canonO > OcleftO > ScleftO > canonS $\geq$ OcleftS $\geq$ ScleftS

(Note that in both cases $\geq$ indicates that object clefts with prosodic prominence on the subject (OcleftS) were not significantly different from either canonical sentences
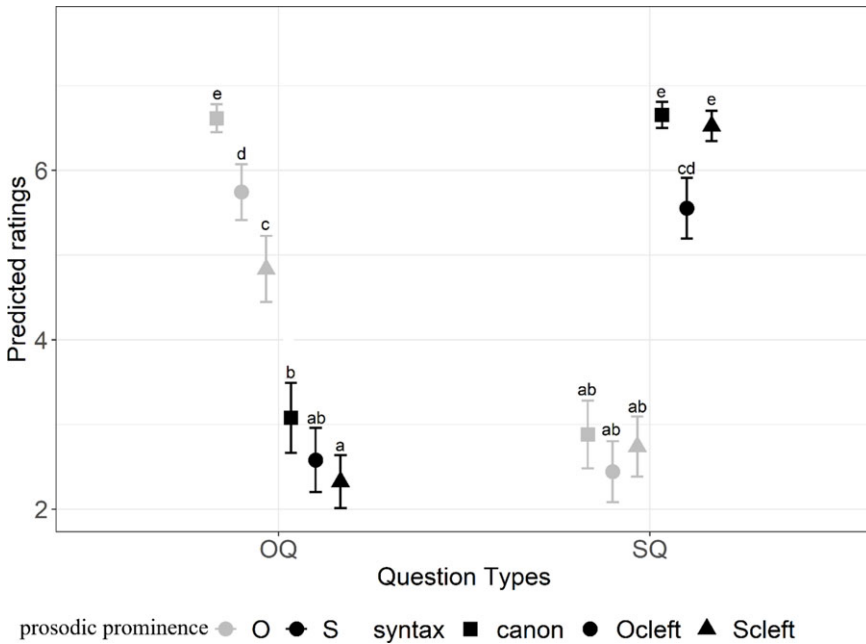
Figure 6. The predicted ratings from cumulative link mixed model for each condition for L1. Gray and black indicate the location of prosodic prominence, and different shapes indicate the different syntactic types.

with prosodic prominence on the subject (canonS) or subject clefts with prosodic prominence on the subject (ScleftS), but canonS and ScleftS were significantly different from each other).

As Figure 6 shows, sentences with prosodic prominence on the intended focus were all rated relatively high (>4) and sentences with no prominence on the intended focus were all relatively low (<4) regardless of whether there was a clefting cue and whether it was consistent or not. This clearly shows that prosodic prominence is an effective focus cue in Mandarin for all three syntactic types (canonical, subject clefts, and object clefts).

With prominence on the focus (grey for OQ, black for SQ in the figure), adding clefting cues, whether consistent or inconsistent, is generally dispreferred (compare circles and triangles with squares): consistent clefting cues resulted in lower ratings for object clefts; inconsistent clefting cues resulted in lower ratings for both subject and object clefts. With no prominence on the focus (black for OQ, grey for SQ), adding clefting cues, whether consistent or inconsistent, mostly played little role: when the question was about the subject and prominence was on the object, ratings for object and subject clefts were not significantly different from canonical sentences. When the question was about the object and prominence was on the subject, subject clefts, but not object clefts, differed significantly from canonical sentences, while the cleft types did not differ from each other.

It is possible that the ratings of sentence types differ from one another regardless of focus marking due to their naturalness or acceptability (compare, for instance,

Table 3. ANOVA table of the final model for ratings for L2

	Chisq	df	p
Syntax	59	2	<.001
ProsodicProminence	2	1	.1
QuestionType	17	1	<.001
Proficiency	5	1	.03
Syntax:ProsodicProminence	51	2	<.001
Syntax:QuestionType	279	2	<.001
ProsodicProminence:QuestionType	609	1	<.001
Syntax:ProsodicProminence:QuestionType	15	2	<.001

Note. Model: Syntax*ProsodicProminence*QuestionType + Proficiency + (1|Item) + (1|Participant).

OcleftO and OcleftS). The results showed that under the same sentence condition, changing the question type resulted in a significant change of ratings for all sentence types. For all sentence types, ratings were higher when the prosodic prominence in the sentence matched the intended focus regardless of whether clefting cues were consistent. This shows the importance of prosodic prominence as focus cues in L1 Mandarin.

L2 interpretation of focus
Table 3 presents the significance of variables in the final model for L2 (the results of the fixed effects are summarized in the On-line Supplementary Materials).

Participants with lower proficiency tend to give higher ratings ($\beta = -0.2$, $SD = 0.1$, $z = -2$). The predicted ratings were calculated from the model and plotted using *ggplot* in Figure 7 with CLD (see the estimates, standard errors, *z*, and *p* values in the On-line Supplementary Materials). The range of L2 listeners' ratings was generally comparable with those for the L1 participants.

Figure 7 shows the following preferences of each answer type for SQ and OQ in the L2 experiment:

(12) SQ: canonS = ScleftS > OcleftS = canonO = ScleftO > OcleftO
      OQ: OcleftO = canonO > ScleftO ≥ OcleftS = canonS ≥ ScleftS

(Note that in both cases ≥ indicates that object clefts with prosodic prominence on the subject (OcleftS) were not significantly different from subject clefts with prosodic prominence on the object (ScleftO), and that canonical sentences with prosodic prominence on the subject (canonS) or subject clefts with prosodic prominence on the subject (ScleftS), but ScleftO and ScleftS were significantly different from each other).

As Figure 7 shows, the sentences with prosodic prominence on the focus and consistent clefting cues were rated relatively high (conditions labeled with *ef* and *f*, ratings ≥ 6). Sentences with no prominence on the focus or inconsistent clefting
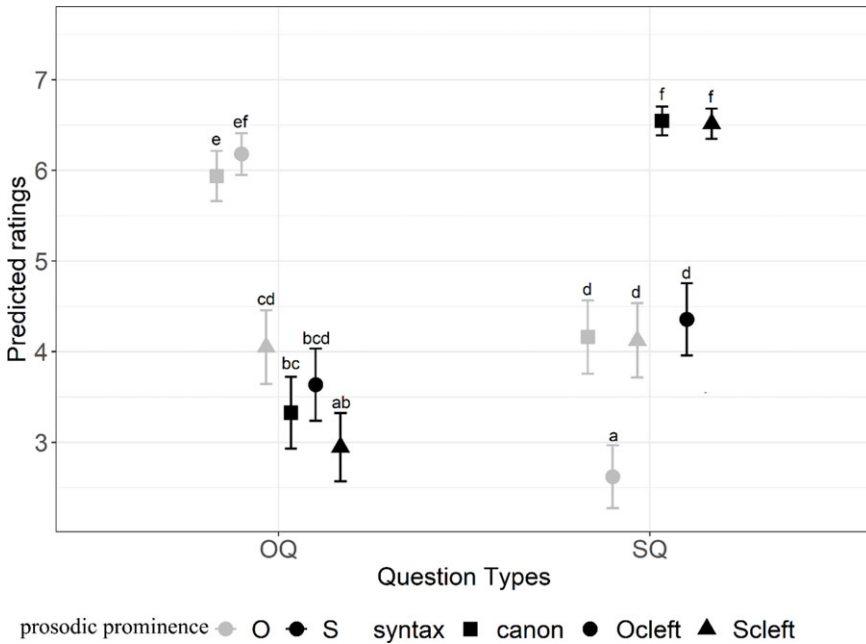
Figure 7. The predicted ratings from cumulative link mixed model for each condition for L2. Gray and black indicate the location of prosodic prominence, and different shapes indicate the different syntactic types.

were judged around the midpoint of the rating scale 4 or lower. This pattern is different from that for Mandarin (see Figure 6) and shows that while prosodic prominence was the main factor in the L1, both prosodic prominence and clefting are important cues to focus for L2 learners.

With prominence on the focus (grey for OQ, black for SQ in the figure), consistent clefting resulted in statistically the same ratings as were obtained for canonical sentences, but inconsistent clefting resulted in lower ratings. With no prominence on the focus (black for OQ, grey for SQ), consistent clefting cues again resulted in statistically the same ratings as for canonical sentences, but now inconsistent clefting cues only resulted in lower ratings if they indicated focus on the object (the first two datapoints for SQ in Figure 7, labeled with *d* and *a*).

When prosodic prominence did not match clefting cues (the "mismatch" sentences ScleftO and OcleftS), the ratings were not significantly different from each other, neither within nor between question types (they all share the letter *d* and were all clustered around the midpoint of the rating scale).

Under the same sentence condition, changing the question type resulted in a significant change of ratings for the canonical sentences and head-stressed clefts, but not for the "mismatch" sentences. The similar ratings to the "mismatch" sentences may suggest that Mandarin learners of English assign equal weighting of prosodic prominence and clefting as cues to focus in English.

Discussion

During speech processing, listeners use various cues to search for important information as part of achieving successful communication. We investigated how L1 Mandarin L2 English listeners of intermediate-to-high proficiency use focus cues in English that are weighed similarly (prosodic cues) and differently (clefting cues) from their L1. We have shown that the role of prosodic prominence as a cue to focus is paramount in L1, but that prosodic and clefting cues are weighted equally in L2. The study suggests that the weighting of cues to focus varies across L1 and L2, contributing to the small body of literature showing cross-linguistic differences in this area. We elaborate on the role of prosodic and clefting cues in L1 and L2 below.

Effect of prosodic prominence

As expected, prosodic prominence was used consistently by Mandarin listeners in rating how appropriate an answer sounds to a question in their L1 and in L2 English. This is consistent with what Calhoun et al. (2021) found for L1 English speakers. This is because prosodic prominence plays an important role in both Mandarin and English and is in line with the predictions in section "Predictions", based on LILt (see section "L2 learning theories"). This suggests that although the phonetic implementation of prosodic focus in the two languages is different, it operates as a functionally equivalent resource of "prosodic prominence", that is, abstract "prominence" transfers straightforwardly from L1 to L2, at least in perception.

As discussed in section "The role of prosody and clefting in the processing of focus in L1 and L2", in reaction time tasks prosodic cues interact with positional cues, for example, sentence-final objects show processing advantages associated with focus regardless of whether they are prosodically prominent. We did not find any such additional effect of position for objects when focus was marked by prosodic prominence in L1 or L2. In L1, the ratings for canonical word order sentences with prominence on the focus were almost identical (near 7), that is, canonS—SQ and canonO—OQ. In L2, the rating for canonS—SQ (*f*) was even higher than canonO—OQ (*e*), the opposite of the expected effect of default focus position. It may be, though, that positional effects would be found in canonical sentences with normal rather than emphatic stress on the object.

However, for L2 listeners, when focus was not marked with prosodic prominence, we found an interesting positional effect that was different to that found in previous L1 studies. For SQs, the ratings for ScleftO and canonO were around the midpoint of the rating scale (= 4). But in their L1, the corresponding ratings were very low, as shown in Figure 8 (note this comparison is indirect as different participants and materials were used in the two experiments). The same was not true for OQs (OcleftS and canonS under OQ), which were rated relatively low in both languages. Note that this effect is different from the positional bias found for L1 listeners in previous studies, where objects were processed as focused even if they were not prominent (equivalent to canonS under OQ). This suggests that L2 listeners are more accepting of "default" prosody in responses (i.e., prominence on
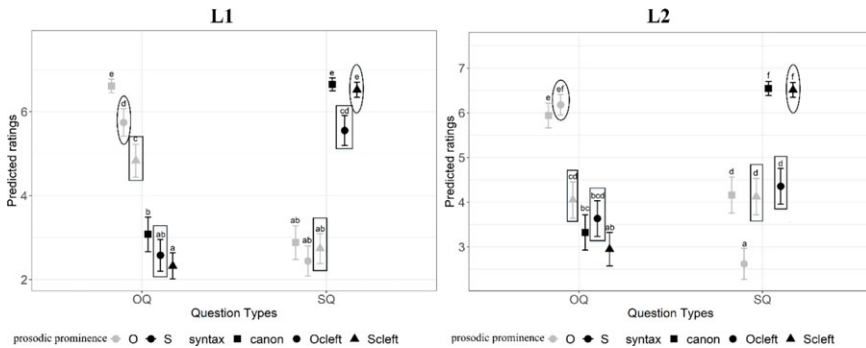
Figure 8. The predicted ratings for each condition in L1 (left panel: Figure 6) and L2 (right panel: Figure 7). Gray and black indicate the prosodic prominence, and shapes indicate the syntactic types. For rectangles and ellipses, see the text.

the object), even where it is not appropriate in the discourse context (i.e., a QUD about the subject).

As mentioned in section "L2 studies", prosody is not normally taught explicitly in classrooms. Although it seems that Mandarin speakers were able to make use of prosodic cues to focus in the current listening study, problems still exist in production and in other aspects of intonation learning (see section "L2 learning theories" on the second dimension of LILt). For example, Mandarin speakers face some difficulties in disambiguating "neutral' broad focus and sentence-final object focus in English in production (Wayland et al., 2019), which likely involves learning differences in phonetic realization. Therefore, some targeted intonation teaching would help learners better produce and comprehend L2 speech. Such teaching is likely to be useful for all second language learners as both closely related and unrelated languages may differ in their mapping between focus and prosody (e.g., German-Dutch in Braun & Tagliapietra, 2011; English-Cantonese in Ge et al., 2020).

Effect of clefting

In relation to the role of clefting and the relative roles of prosodic and clefting cues to focus, the results showed that although clefting cues played a role in both L1 and L2, this was generally inhibitory or neutral rather than facilitatory. In both the L1 and L2, when prosodic and clefting cues to focus were consistent with one another and with the question type (ScleftS for SQ and OcleftO for OQ, indicated in ellipses in Figure 8) the answers were given high ratings. However, consistent clefting did not result in higher ratings than canonical sentences with prominence on the focus, for example, canonS and ScleftS were rated equally highly to the SQ. This is consistent with Arnhold's (2021) rating study on L1 English and also Calhoun et al. (2021), based on question preferences. Based on the CM, when the two cues converged on the same word, the cue strength is strongest, and the answer was therefore rated as highly acceptable (approximately 7) in both their L1 and L2, compared to when the cues diverged (i.e., ScleftO and OcleftS). However, the

combination of the two cues is not any stronger than the prosodic cue alone. It is possible there may be a ceiling effect that obscures any "benefit" that the additional clefting cue might add.

For object focus in Mandarin, OcleftO sentences were rated slightly lower than the corresponding canonical sentences (canonO). This may be because the relative rarity of object *SHI...DE* clefts might make OcleftO somewhat less acceptable, compared to canonO (see section "Syntactic focus marking (clefting)"). However, the ratings for OQ-OcleftO were still high (around 6), indicating that object clefts were still broadly acceptable for native listeners, and that they were more acceptable than the "mismatch" sentences (e.g., OQ-ScleftO) (<3), as shown in Figure 6. In English, having a consistent clefting cue did not lower the appropriateness rating (canonO = OcleftO) (>6), showing differences between L1 and L2.

When the prosodic and clefting cues clashed (the "mismatch" sentences indicated in rectangles in Figure 8), we saw a different relative cue-weighting of prosodic and clefting cues in L1 and L2. Prosodic cues were clearly favored over clefting cues for L1 Mandarin listeners (e.g., SQ > OQ for OcleftS; OQ > SQ for ScleftO). This indicates that prosodic prominence is more important as a cue to focus than clefting in Mandarin. Note that the current study only investigated *SHI...DE* clefts. This does not necessarily mean that all syntactic focus cues are less effective. Pseudo-clefts are more commonly used structures to mark object focus in Mandarin, and they have a dedicated sentence-final focus position for both subject and object clefts. It would be valuable to investigate pseudo-clefts in future studies to see if they play a stronger role than *SHI...DE* in the identification of focus due to the salient position of focus.

In L2, there was no significant difference in the ratings of the "mismatch" sentences (ScleftO and OcleftS) for either question type (i.e., they all share letter *d* in the right-hand panel of Figure 8). This is different to Calhoun et al. (2021), which showed English listeners weighted cleft cues more highly than prosodic cues in such structures to perceive focus in response to a *wh*-question (e.g., participants chose SQs as appropriate for ScleftO sentences). Listeners were therefore weighting clefting cues more strongly in L2 than in L1, but still not in a native-like way. This is consistent with LILt, which predicts difficulties where there are differences in the "semantic dimension," in this case the relative importance of clefting in Mandarin and English focus. While the model did not specifically discuss the inter-action with other linguistic systems, we have shown that applying the theory in this way successfully predicted the difficulty. In addition, based on the CM which predicts that L2 learners need to acquire new patterns of cue-strengths, we could see evidence of learning, in that the Mandarin listeners had learnt that clefting cues are weighted more highly in English (when the two cues clash) than in their L1. However, the rela-tive difficulty of integrating these cues (or resolving the competition between them) ultimately made these structures ambiguous as to whether they marked subject or object focus for these L2 listeners. These results have shown that there is a negative transfer to L2 of cue strengths that derive from Mandarin listeners' L1 pattern of cue strengths (i.e., the fact that in their L1 the prosodic cue wins the competition when prosodic and clefting cues place focus on different words in an utterance). The CM is not specifically directed toward L2 learning of cue competition between prosody and clefting, but the theory can be applied in this way.

In this paper, we have looked at learners whose L1 is Mandarin, where clefting seems to be a less important marker of focus than prosodic prominence. It would be valuable to look at learners with other L1s, especially those in which clefting is as or more important than prosodic cues to focus such as Korean (Kember et al., 2021), or where prosodic prominence is an optional marker of focus such as Yucatec Maya (Kügler & Skopeteas, 2006). To our knowledge, there are no such studies; however, we predict that for L2 English learners with such L1s, syntactic focus marking cues will be more important than prosodic in the L2, and they will find it relatively harder to make effective use of prosodic cues (similar to the findings of Ge et al. 2020). In addition, it would be particularly interesting to look at L2 learners with L1s like Russian, where prosodic and syntactic cues seem to be equally important in the perception of prominence, and these cues interact with each other in a complex way (Luchkina & Cole, 2021). Regarding LILt, such learners should find it relatively straight-forward to effectively use clefting and prominence cues to focus in English, since they both have that semantic function in their L1. However, regarding the CM, difficulties may arise in acquiring the particular patterns of cue strength that differ between their L1 and L2. In general, we predict that the importance of different focus markers in the L1 will be linked to learners' performance in their L2, rather than particular types of focus marker (prominence, clefts, etc.) being inherently more or less difficult for L2 learners to acquire.

In this study, we recruited learners in a relatively narrow proficiency range: inter-mediate-to-high. As expected, therefore, while we included proficiency as a control factor in the L2 experiment model, it did not interact with prosodic prominence, syntax or question type; although we did find a general effect of proficiency, in that participants with lower proficiency tend to give higher ratings. In future research, it would be interesting to look at the learning trajectory by including learners from a wider range of proficiencies, as L2 proficiency is an important factor in the compre-hension of focus in L2 (Akker & Cutler, 2003; Braun & Tagliapietra, 2011; Lee & Fraundorf, 2017). Previous L2 studies based on the CM have predicted that L2 profi-ciency is an influential factor which determines the developmental direction of cue use as well as transfer patterns (differentiation, forward transfer, backward transfer, and amalgamation, see Li & MacWhinney, 2013). Future studies could investigate learners with different levels of L2 proficiency to shed light on the acquisition of cue-focus mappings.

Conclusion

This paper has presented an investigation of the role of prosodic and clefting cues to focus by Mandarin listeners in their L1 and L2. We have shown that prosodic prom-inence cues to focus are used in a functionally equivalent way in L1 and L2, similar to native speakers in both languages. However, L2 listeners found it more difficult to integrate the relatively higher-weighted clefting cues in English. In future work, this method could be extended to different prosodic conditions (e.g., "neutral" broad focus), to different clefting types (e.g., pseudo-clefts), and to learners with different levels of L2 proficiency. It would also be good in future work to test the relative importance and use of different cues to focus for learners from a wider variety

of L1s and to increase our understanding of how use of focus cues in the L1 and L2 relate to each other.

Conflict of interest. The author(s) declare none.

Notes

1. Please note that "I wonder what the fox is licking" is not actually a question, although pragmatically it functions like a question. We follow the labels from Ge et al. (2020).
2. Our study focused on learners with a similar level of proficiency (moderate to high proficiency), so we did not predict any effect of proficiency. Nonetheless, there would inevitably be small differences in proficiency between learners, so this was included as a control factor.

References

Akker, E., & Cutler, A. (2003). Prosodic cues to semantic structure in native and nonnative listening. *Bilingualism: Language and Cognition*, **6**(2), 81–96. https://doi.org/10.1017/S1366728903001056

Arnhold, A. (2021). Prosodic focus marking in clefts and syntactically unmarked equivalents: Prosody – syntax trade-off or additive effects? *Journal of the Acoustical Society of America*, **149**(3), 1390–1399. https://doi.org/10.1121/10.0003594

Bates, E., & MacWhinney, B. (1989). Functionalism and the competition model. In B. MacWhinney & E. Bates (Eds.), *The crosslinguistic study of sentence processing* (pp. 3–73). Cambridge University Press.

Birch, S. L., & Clifton, C. (1995). Focus, accent, and argument structure: Effects on language comprehension. *Language and Speech*, **38**(4), 365–391. https://doi.org/10.1177/002383099503800403

Boersma, P., & Weenink, D. (2018). Praat: doing phonetics by computer (Version 6.0.37) [Computer software]. http://www.praat.org

Braun, B., & Tagliapietra, L. (2011). On-line interpretation of intonational meaning in L2. *Language and Cognitive Processes*, **26**(2), 224–235. https://doi.org/10.1080/01690965.2010.486209

Breen, M., Fedorenko, E., Wagner, M., & Gibson, E. (2010). Acoustic correlates of information structure. *Language and Cognitive Processes*, **25**(7-9), 1044–1098. https://doi.org/10.1080/01690965.2010.504378

Brugos, A., Shattuck-Hufnagel, S., & Veilleux, N. (2006). Transcribing prosodic structure of spoken utterances with ToBI. https://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-911-transcribing-prosodic-structure-of-spoken-utterances-withtobijanuary-iap-2006/index.htm

Calhoun, S. (2010). The centrality of metrical structure in signaling information structure: A probabilistic perspective. *Language*, **86**(1), 1–42. https://doi.org/10.1353/lan.0.0197

Calhoun, S., Wollum, E., & Kruse-Va'ai, E. (2021). Prosodic prominence and focus: Expectation affects interpretation in Samoan and English. *Language and Speech*, **64**(2), 346–380. https://doi.org/10.1177/0023830919890362

Carlson, K., Dickey, M. W., Frazier, L., & Clifton, C. (2009). Information structure expectations in sentence comprehension. *The Quarterly Journal of Experimental Psychology*, **62**(1), 114–139. https://doi.org/10.1080/17470210701880171

Chafe, W. (1976). Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In C. N. Li (Ed.), *Subject and topic* (pp. 25–55). New York: Academic Press.

Chao, Y. R. (1947). *Cantonese primer*. Harvard University Press.

Chao, Y. R. (1968). *A grammar of spoken Chinese*. University of California Press.

Chen, A., & Lai, V. (2011). Comb or coat: The role of intonation in online reference resolution in a second language. In W. Zonneveld, H. Quené, & W. Heeren (Eds.), *Sound and Sounds. Studies presented to M.E.H. (Bert) Schouten on the occasion of his 65th birthday* (pp. 57–68). UiL OTS.

Chen, S.-H., Chen, S.-C., & He, T.-H. (2012). Surface cues and pragmatic interpretation of given/new in Mandarin Chinese and English: A comparative study. *Journal of Pragmatics*, **44**(4), 490–507. https://doi.org/10.1016/j.pragma.2011.12.006

Chen, Y., & Braun, B. (2006). Prosodic Realization of Information, Structure Categories in Standard Chinese. In R. Hoffmann, H. Mixdorff (Eds.), *Speech prosody 2006* (p. 54). Dresden, Germany: TUD Press.

Christensen, R. (2015). Ordinal – Regression models for ordinal data (R package version 2019.3-9).

Clifton, C., & Frazier, L. (2016). Focus in corrective exchanges: Effects of pitch accent and syntactic form. *Language and Speech*, **59**(4), 544–561. https://doi.org/10.1177/0023830915623578

Cobb, T. (2000). *The Compleat lexical tutor [Website]*. https://www.lextutor.ca/vp/comp/

Delin, J., & Oberlander, J. (1995). Syntactic constraints on discourse structure: The case of it-clefts. *Linguistics*, **33**(3), 465–500. https://doi.org/10.1515/ling.1995.33.3.465

Destruel, E., Velleman, D., Onea, E., Bumford, D., Xue, J., & Beaver, D. (2013). A crosslinguistic study of the non-at-issueness of exhaustive inferences. In F. Schwarz (Ed.), *Experimental perspectives on presup-positions* (pp. 135–156). Switzerland: Springer. DOI: https://doi.org/10.1007/978-3-319-07980-6_6

É Kiss, K. (1998). Identificational focus versus information focus. *Language*, **74**(2), 245–273. https://doi.org/10.2307/417867

Féry, C., & Ishihara, S. (Eds.) (2016). *The Oxford handbook of information structure*. Oxford University Press.

Foltz, A. (2020). Using prosody to predict upcoming referents in the L1 and the L2: The role of recent exposure. *Studies in Second Language Acquisition*, 1–28. https://doi.org/10.1017/S0272263120000509

Ge, H., Chen, A., & Yip, V. (2020). Comprehension of focus-to-accentuation mapping in sentences with only by advanced Cantonese learners and Dutch learners of English. *Studies in Second Language Acquisition*, 1–25. https://doi.org/10.1017/S0272263120000248

Graves, S., Piepho, H.P., & Selzer, L. (2019). multcompView: Visualizations of paired comparisons. (Version 0.1-8). https://cran.rproject.org/web/packages/multcompView/index.html

Greif, M. (2010). Contrastive focus in Mandarin Chinese. *Proceedings of speech prosody 2010* (pp. 1–4), Chicago

Greif, M., & Skopeteas, S. (2021). Correction by Focus: Cleft Constructions and the Cross-Linguistic Variation in Phonological Form. *Frontiers in Psychology*, **12**, 648478.

Hedberg, N. (1999). The discourse function of English clefts and Mandarin Shi... de constructions. In *Workshop on the discourse function of clefts* (pp. 1–3). Humboldt University, Berlin, Germany.

Hedberg, N. (2013). Multiple focus and cleft sentences. In K. Hartmann & T. Veenstra (Eds.), *Cleft structures* (pp. 227–250). John Benjamins.

Hedberg, N., & Sosa, J. M. (2008). The prosody of topic and focus in spontaneous English dialogue. In C. Lee, M. Gordon & D. Büring (Eds.), *Topic and focus: Cross-linguistic perspectives on meaning and intonation* (pp. 101–120). Dordrecht: Springer.

Hervé, M. (2015). RVAideMemoire: diverse basic statistical and graphical functions (R package version 0.9-50).

Hole, D. (2011). The deconstruction of Chinese shì...de clefts revisited. *Lingua*, **121**(11), 1707–1733. https://doi.org/10.1016/j.lingua.2011.07.004

Hole, D. (2012). The information structure of Chinese. In M. Krifka & R. Musan (Eds.), *The expression of information structure* (pp. 45–70). De Gruyter Mouton.

Ip, M. H. K., & Cutler, A. (2020). Universals of listening: Equivalent prosodic entrainment in tone and non-tone languages. *Cognition*, **202**. https://doi.org/10.1016/j.cognition.2020.104311

Jenkins, J. (2002). A sociolinguistically based, empirically researched pronunciation syllabus for English as an international language. *Applied Linguistics*, **23**, 83–103. https://doi.org/10.1093/applin/23.1.83

Kember, H., Choi, J., Yu, J., & Cutler, A. (2021). The processing of linguistic prominence. *Language and Speech*, **64**(2), 413–436. https://doi.org/10.1177/0023830919880217

Kisler, T., Reichel, U., & Schiel, F. (2017). Multilingual processing of speech via web services. *Computer Speech and Language*, **45**, 326–347. https://doi.org/10.1016/j.csl.2017.01.005

Krifka, M. (2008). Basic notions of information structure. *Acta Linguistica Hungarica*, **55**, 243–276.

Kügler, F., & Calhoun, S. (2020). Prosodic encoding of information structure: A typological perspective. In C. Gussenhoven & A. Chen (Eds.), *The Oxford handbook of language prosody*. Oxford University Press.

Kügler, F., & Skopeteas, S. (2006). Interaction of Lexical Tone and Information Structure in Yucatec Maya. *Proceedings of the Second International Symposium on Tonal Aspects of Languages* (pp. 77–82). La Rochelle.

Ladd, D. R. (1996/2008). *Intonational phonology*. Cambridge: Cambridge University Press.

Ladd, D. R., & Schepman, A. (2003). "Sagging transitions" between high pitch accents in English: experimental evidence. *Journal of Phonetics*, **31**(1), 81–112. https://doi.org/10.1016/S0095-4470(02)00073-6

Lambrecht, K. (2001). A framework for the analysis of cleft constructions. *Linguistics*, **39**(3), 463–516. https://doi.org/10.1515/ling.2001.021

Lee, E.-K. & Fraundorf, S. H. (2017). Effects of contrastive accents in memory for L2 discourse. *Bilingualism: Language and Cognition*, **20**(5), 1063–1079. https://doi.org/10.1017/S1366728916000638

Lee, E.-K., & Fraundorf, S. H. (2019). Native-like processing of prominence cues in L2 written discourse comprehension: Evidence from font emphasis. *Applied Psycholinguistics*, **40**(2), 373–398. https://doi.org/10.1017/S0142716418000619

Lee, Y.-C., Wang, T., & Liberman, M. (2016). Production and Perception of Tone 3 Focus in Mandarin Chinese. *Frontiers in Psychology*, **7**, 1058. https://doi.org/10.3389/fpsyg.2016.01058

Li, P., and MacWhinney, B. (2013). Competition model. In C. A. Chapelle (Ed.), *The Encyclopedia of Applied Linguistics* (pp. 1–5). Malden, MA: John Wiley and Sons, Inc.

Liu, H., Bates, E., & Li, P. (1992). Sentence interpretation in bilingual speakers of English and Chinese. *Applied Psycholinguistics*, **13**, 451–84.

Liu, W., & Kempson, R. (2018). Chinese cleft structures and the dynamics of processing. *Transactions of the Philological Society*, **116**(1), 91–116. https://doi.org/10.1111/1467-968X.12106

Liu, Y., & Yang, Y. (2016). Exhaustivity in Mandarin Shi . . . (de) sentences: experimental evidence. In M. Köllner & R. Ziai (Eds.), *Proceedings of European summer school in logic language and information 2016 student session* (pp. 167–179). Bolzano: Free University of Bozen.

Luchkina, T., & Cole, J. (2021). Perception of word-level prominence in free word order language discourse. *Language and Speech*, **64**(2), 381–412. https://doi.org/10.1177/0023830919884089

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, **44**(2), 314–324. https://doi.org/10.3758/s13428-011-0168-7

McAuliffe, M., Socolof, M., Mihuc, S., Wagner, M., & Sonderegger, M. (2017). Montreal Forced Aligner (Version 0.9.0) [Computer software].

Mennen, I. (2015). Beyond segments: Towards a L2 intonation learning theory. In Elisabeth Delais-Roussairie, Mathieu Avanzi, and Sophie Herment (Eds.), *Prosody and language in contact* (pp. 171–188). Springer.

Molnár, V. (2006). On different kinds of contrast. In V. Molnár & S. Winkler (Eds.), *The architecture of focus* (pp. 197–233). Mouton de Gruyter.

Onea, E. (2019). Exhaustivity in it-clefts. In C. Cummins & N. Katsos (Eds.), *The Oxford handbook of experimental semantics and pragmatics*. Oxford, UK: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198791768.013.17

Ortega-Llebaria, M. and Colantoni, L. (2014). L2 English intonation: Relations between form-meaning associations, access to meaning, and L1 transfer. *Studies in Second Language Acquisition* **36**(2), 331–353. https://doi.org/10.1017/S0272263114000011

Paul, W., & Whitman, J. (2008). Shi . . . de Focus Clefts in Mandarin Chinese. *Linguistic Review*, **25**(3-4), 413–451. https://doi.org/10.1515/TLIR.2008.012

Pierrehumbert, J., & Hirschberg, J. (1990). The meaning of intonational contours in the interpretation of discourse. In P. Cohen, J. Morgan & M. Pollack (Eds.), *Intentions in communication* (pp. 271–311). Cambridge MA: MIT Press.

Prince, E. F. (1978). A comparison of wh-clefts and it-clefts in discourse. *Language*, **54**(4), 883–906. https://doi.org/10.2307/413238

R Core Team. (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing: Vienna, Austria. https://www.R-project.org/

Reichle, R., & Birdsong, D. (2014). Processing focus structure in L1 and L2 French: L2 proficiency effects on ERPs. *Studies in Second Language Acquisition*, **36**(3), 535–564. doi: 10.1017/S0272263113000594

Roberts, C. (1996). Information structure in discourse: Towards an integrated formal theory of pragmatics. In J. H. Yoon & A. Kathol (Eds.), *OSU Working Papers in Linguistics 49: Papers in Semantics* (pp. 91–136). Columbus: The Ohio State University.

Roland, D., Dick, F., & Elmanc, J. (2007). Frequency of basic English grammatical structures: A corpus analysis. *Journal of Memory and Language*, **57**(3), 348–379. https://doi.org/10.1016/j.jml.2007.03.002

Rooth, M. (1992). A theory of focus interpretation. *Natural Language Semantics*, **1**(1), 75–116. https://doi.org/10.1007/BF02342617

Teng, S.-H. (1979). Remarks on cleft sentences in Chinese. *Journal of Chinese Linguistics*, **7**(1), 101–114.

Trouvain, J., Gut, U., & Barry, W. J. (2007). Bridging research on phonetic descriptions with knowledge from teaching practice: The case of prosody in non-native speech. In J. Trouvain & U. Gut (Eds.), *Non-native prosody: Phonetic description and teaching experience* (pp. 3–21). Mouton de Gruyter.

Vallduví, E. (2016). Information structure. In M. Aloni & P. Dekker (Eds.), *The Cambridge handbook of formal semantics* (pp. 728–755). UK: Cambridge University Press. https://doi.org/10.1017/CBO9781139236157.024

van Maastricht, L., Krahmer, E., Swerts, M. (2016). Prominence patterns in a second language: Intonational transfer from Dutch to Spanish and vice versa. *Language Learning*, **66**(1), 124–158.

Wayland, R., Guerra, C., Chen, S., & Zhu, Y. (2019). English focus perception by Mandarin listeners. *Languages*, **4**(4), 91. https://doi.org/10.3390/languages4040091

Welby, P. (2003). Effects of pitch accent position, type, and status on focus projection. *Language and Speech*, **46**(1), 53–81. https://doi.org/10.1177/00238309030460010401

Wickham, H., Chang, W., & Wickham, M. H. (2016). Package 'ggplot2'. Create elegant data visualisations using the grammar of graphics (Version 3.0.0).

Xu, L. (2004). Manifestation of informational focus. *Lingua*, **114**(3), 277–299. https://doi.org/10.1016/S0024-3841(03)00031-7

Xu, Y. (1999). Effects of tone and focus on the formation and alignment of F0 contours. *Journal of Phonetics*, **27**(1), 55–105. https://doi.org/10.1006/jpho.1999.0086

Yan, M., & Calhoun, S. (2019). Priming effects of focus in Mandarin Chinese. *Frontiers in Psychology*, **10**, 1985. https://doi.org/10.3389/fpsyg.2019.01985

Yan, M., & Calhoun, S. (2020). Rejecting false alternatives in Chinese and English: the interaction of prosody, syntax and default focus position. *Laboratory Phonology*, **11**(1). https://doi.org/10.5334/labphon.255

Yan, M., Calhoun, S., & Warren, P. (2022). The role of prominence in activating focused words and their alternatives in Mandarin: Evidence from lexical priming and recognition memory. *Language and Speech*. https://doi.org/10.1177/00238309221126108