Journal of Memory and Language 67 (2012) 149–164

Contents lists available at SciVerse ScienceDirect

Journal of Memory and Language

journal homepage: www.elsevier.com/locate/jml

Sometimes children are as good as adults: The pragmatic use of prosody in children's on-line sentence processing

Peng Zhou *, Stephen Crain, Likan Zhan

Macquarie University, Sydney, NSW 2109, Australia

ARTICLE INFO

Article history:
Received 2 November 2011
revision received 16 March 2012
Available online 12 April 2012

Keywords:
Sentence processing
Prosody
Speech act
Mandarin Chinese
Child language

ABSTRACT

This study examined 4-year-old Mandarin-speaking children's sensitivity to prosodic cues in resolving speech act ambiguities, using eye-movement recordings. Most previous on-line studies have focused on children's use of prosody in resolving structural ambiguities. Although children have been found to be sensitive to prosodic information, they use such information less effectively than adults in on-line sentence processing. The present study takes advantage of special properties of Mandarin Chinese to investigate the role of prosody in children's on-line processing of ambiguities in which prosody serves to signal the illocutionary meaning of an utterance (i.e., whether the speaker is asking a question or making a statement). We found that the effect of prosody in this case was as robust in children as it was in adults. This suggests that children are as sensitive as adults in using prosody in on-line sentence processing, when prosody is used to resolve a pragmatic ambiguity.

Introduction

When presented with spoken sentences that contain ambiguities of certain kinds, adults are able to use prosodic information to immediately arrive at the intended interpretation (e.g., Ferreira, Anes, & Horine, 1996; Marslen-Wilson, Tyler, Warren, Grenier, & Lee, 1992; Nagel, Shapiro, Tuller, & Nawy, 1996; Snedeker & Trueswell, 2003; Snedeker & Yuan, 2008; Speer, Kjelgaard, & Dobroth, 1996; Warren, Grabe, & Nolan, 1995; Weber, Grice, & Crocker, 2006). However, mastery of the interface of prosodic information and syntactic structure takes time, as it has been found that children younger than six or seven are less adept than adults at using prosodic cues to determine the intended meaning of ambiguous sentences. This difference between children and adults is a robust finding, covering several types of ambiguities, and several languages (Choi & Mazuka, 2003; Ito, Jincho, Minai, Yamane, & Mazuka, 2012; Sekerina & Trueswell, in press; Snedeker & Yuan, 2008; Zhou, Su, Crain, Gao, & Zhan, in press).

The present study investigates a new kind of ambiguity in which prosody plays a critical role. The ambiguity involves speech acts, whether someone is asking a question or making a statement. In Mandarin Chinese, the same sequence of words can be used to make either speech act; which speech act is being performed can be conveyed by prosodic information. There are noteworthy differences between this kind of 'pragmatic' ambiguity and the kinds of ambiguities that have been investigated previously, where children have been found to be less sensitive than adults in using prosody to resolve the ambiguity. To anticipate the experimental findings, Mandarin-speaking children proved to be the equals of adults in using prosodic cues in resolving such pragmatic ambiguities in on-line sentence processing. Before we introduce our study of this novel type of ambiguity, we briefly review the findings from previous research on children's use of prosody to resolve other kinds of ambiguities.

Charting the course of children's development of the mapping between prosody and other levels of linguistic knowledge (e.g., syntax, semantics and pragmatics) has received considerable attention in the literature on child

* Corresponding author. Address: ARC Centre of Excellence in Cognition and its Disorders, Macquarie University, Sydney, NSW 2109, Australia. Fax: +61 2 9850 6059.

E-mail address: peng.zhou@mq.edu.au (P. Zhou).

0749-596X/$ - see front matter
http://dx.doi.org/10.1016/j.jml.2012.03.005

sentence processing. Two main questions have been asked: (i) what is the nature of the differences between the use of prosodic information by adults and by children? (ii) To what extent are the problems experienced by children due to the demands made by on-line sentence processing? Previous research on children's use of prosodic information in sentence processing has focused on three main areas: (a) the role of prosody in establishing phrase boundaries (Choi & Mazuka, 2003; Snedeker & Yuan, 2008), (b) the role of prosody in associating focus particles with appropriate expressions (Höhle, Berger, Müller, Schmitz, & Weissenborn, 2009; Zhou et al., in press) and (c) the role of prosody in interpreting contrastive focus (Ito et al., 2012; Sekerina & Trueswell, in press). Let us review the main findings in each of these areas of research.

In studies of children's use of prosodic cues in establishing phrase boundaries, the findings have been mixed. For example, Choi and Mazuka (2003) tested Korean-speaking children on two types of ambiguities: word-segmentation ambiguities and phrasal grouping ambiguities. They found that 3–4-year-old Korean-speaking children had no difficulty in using prosodic boundary cues to resolve ambiguities at the word-level, i.e., in segmenting words. However, children were unable to use similar prosodic cues to resolve ambiguities at the sentence-level, i.e., to establish phrasal groupings. In another representative study, Snedeker and Yuan (2008) investigated whether English-speaking children can use prosodic boundary cues to disambiguate sentences like 'Tap the frog with the flower'. There are two possible attachments of the prepositional phrase *with the flower* in this sentence. It can be associated either with the verb *tap* or with the noun phrase *the frog*, yielding two possible readings: the instrument reading on which the sentence conveys the instruction to 'tap the frog using the flower', and the modifier reading on which the sentence conveys the instruction to 'tap the frog that has the flower'. Prosodic boundary cues can be used to distinguish between the two interpretations. A prosodic boundary between *frog* and *with* encourages the instrument reading, whereas a prosodic boundary between *tap* and *the* signals the modifier reading. Using eye-movement recordings, Snedeker and Yuan examined whether young English-speaking children can use prosodic boundaries to resolve such ambiguities. Children were presented auditorily with either of the two prosodic versions of the test sentences and were asked to act upon objects in the experimental workspace, based on their interpretation of the sentences. It was found that 4–5-year-old English-speaking children responded differently to the two versions of the spoken sentences; children fixated longer on the intended object based on the prosodic information provided, though this prosodic effect was delayed by approximately 500 ms, as compared to adults. Snedeker and Yuan concluded that children are able to use prosodic boundary cues in resolving ambiguities of phrasal attachment. Note that the prosodic effect elicited in the Snedeker and Yuan study resulted from a blocked design. Half of the child participants heard the instrument version of the sentences in the first session of the experiment, and heard the modifier version in the second session. The other half were presented with the same sentences, but in reverse order (cf. Snedeker & Trueswell, 2001). A blocked design, Snedeker and Yuan reasoned, was more likely to be successful in detecting children's sensitivity to prosody, as compared to having the different attachment possibilities interspersed in the same session. In the Choi and Mazuka (2003) study, which found no reliable effect of prosody on phrase structure attachment, both versions of the spoken sentences (one with a phrase boundary and one without the phrase boundary) were presented in the same experimental session. Snedeker and Yuan (2008) reasoned that this design might have caused a contamination effect, that is, a strong tendency to perseverate across trials could easily wipe out a small or fragile effect of prosody.

Another sentence-level ambiguity can often be resolved using prosodic information. This ambiguity involves the association of focus particles with appropriate expressions elsewhere in a sentence. A positive result in this areas was achieved by Höhle et al. (2009), who investigated how German-speaking children interpret sentences with accented and unaccented focus particle *auch* 'also', using eye-tracking. German sentences like (1) are ambiguous.

(1) Toby hat auch einen Stift.
Toby has also a pen
a. $[\text{Toby}]_F$ hat AUCH einen Stift. (in addition to someone else having a pen)
b. Toby hat auch einen $[\text{STIFT}]_F$. (in addition to having something else)

Sentence (1) can have at least two interpretations depending on where the pitch accent is placed. Throughout the text the focused element is indicated by F-brackets and the pitch accent is indicated by capitals. When a pitch accent is placed on the focus particle *auch* 'also', the focus particle is associated with the subject noun phrase *Toby*, as indicated in (1a). So the sentence entails that in addition to Toby there is someone else who has a pen. When a pitch accent is placed on the object noun phrase *Stift* 'pen', the focus particle *auch* 'also' is associated with the object noun phrase, as indicated in (1b). In this case, the sentence entails that Toby has something else, in addition to a pen. Höhle et al. found that the eye gaze patterns of 3–4-year-old German-speaking children accurately reflected the prosody of the spoken sentences. The conclusion is that German-speaking children use pitch accent to find the correct focused elements associated with focus particles. Another study investigating children's sensitivity to pitch accent in associating focus particles with appropriate expressions is by Zhou et al. (in press). This study investigated whether 4-year-old Mandarin-speaking children use pitch accent to disambiguate sentences like (2). Again, the methodology was eye-tracking.

(2) Zhiyou Yuehan de pingguo shi hongde.
Only John DE apple is red
a. Zhiyou $[\text{Yuehan de PINGGUO}]_F$ shi hongde.
Meaning: John's apple is red, and nothing else is red.
b. Zhiyou $[\text{YUEHAN}]_F$ de pingguo shi hongde.
Meaning: John's apple is red, and no one else's apple is red.

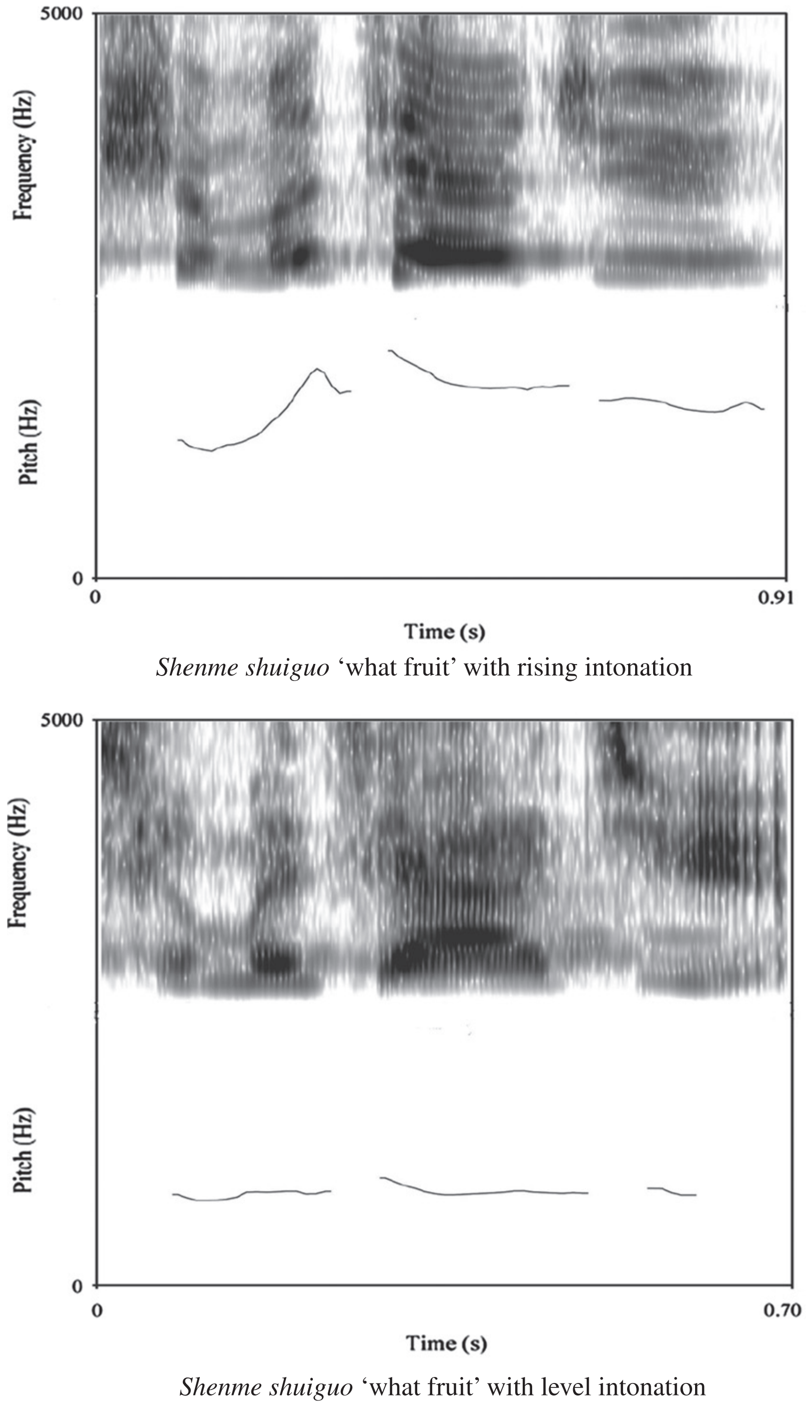


Fig. 1. Intonation contours for the *wh*-phrase *shenme shuiguo* 'what fruit' with rising intonation (upper panel) and with level intonation (lower panel).

Sentence (2) is ambiguous in Mandarin Chinese. The focus particle *zhiyou* 'only' can associate either with the entire subject noun phrase *Yuehan de pingguo* 'John's apple' or with the modifier of the subject noun phrase *Yuehan* 'John'. Pitch accent provides the necessary information to disambiguate between the two associations. A pitch accent on the head noun *pingguo* 'apple' encourages the interpretation in (2a); and a pitch accent on the modifier *Yuehan* 'John' encourages the interpretation in (2b). Zhou et al. found that the fixation patterns of 4-year-old Mandarin-speaking children indicated that they use pitch accent to resolve ambiguities involving the focus particle *zhiyou* 'only', though the effect of prosody was delayed in children, as compared to adults (as in the Snedeker and Yuan study).

Another linguistic phenomenon in which prosody plays a role is in interpreting contrastive focus. Sekerina and

Trueswell (in press) investigated the role of pitch accent in Russian-speaking children's processing of contrastive focus. For example, when adults hear the adjective + noun expression 'red butterfly' with a pitch accent on the adjective, as in *RED butterfly*, they infer the existence of at least one non-red butterfly in the context. By contrast, when a pitch accent is placed on the noun, as in *red BUTTERFLY*, adults infer that the context contains at least one other red entity, in addition to at least one red butterfly. The Sekerina and Trueswell study revealed that 6-year-old Russian-speaking children used pitch accent on the adjective (as in *RED butterfly*) to facilitate identification of the intended referent, but only when the adjective and noun were adjacent (non-split constructions). If a verb intervened between the adjective and the noun (split constructions), an early contrastive pitch accent did not facilitate the identification of the intended referent. This facilitative effect of pitch accent was also found in 6-year-old Japanese-speaking children's processing of contrastive focus (Ito et al., 2012). Note that both Russian-speaking and Japanese-speaking children exhibited delayed facilitative effect of prosody, as compared to adults.

Taken together, the findings of previous research suggest that, although children are able to use prosodic information in on-line sentence processing, they use such information less effectively than adults in deciding on the intended interpretation, as evidenced by the fact that children showed delayed effect of prosody, as compared to adults (Ito et al., 2012; Sekerina & Trueswell, in press; Snedeker & Yuan, 2008; Zhou et al., in press). Prosodic information can be integrated rapidly by adults in arriving at an intended interpretation, but children are less able to do that.

We attribute the difference between children and adults to the particular properties of the grammar that have been investigated in prior research. More specifically, we propose that the delayed effect of prosody found in children is due to the nature of the specific mapping between prosody and syntax in child grammar. During the comprehension of a sentence, listeners are engaged in the recovery of syntactic, semantic and prosodic characterisations of the linguistic input. Each such characterisation is computed within its own representational system, but these representations also interact with one another over time as the sentence unfolds. Most of these on-line studies have focused on the role of prosody in resolving syntactic ambiguities. The resolution of syntactic ambiguities involves the interaction between syntax and prosody. Adults are able to compute and integrate the two levels of linguistic representations rapidly and effortlessly to arrive at an intended interpretation of the linguistic input, due to the well-established mapping between a syntactic representation and its corresponding prosodic representation. Children, however, are less able to do that, since the mapping between syntax and prosody in child language is not as well-established as in adult language. Children might have adult-like representations within each basic level, i.e., they have adult-like syntactic representation of a sentence as well as adult-like prosodic representation of the sentence. Yet, the association between the syntactic representation and its corresponding prosodic representation is less automatic for children than it is for adults when processing a sentence, therefore more processing time is needed for children before they can reach an intended interpretation, yielding a delayed effect of prosody.

Along this line of reasoning, a further question to ask is how other linguistic interfaces are represented in children's sentence processing system. Are different linguistic interfaces equally represented in children's sentence processing system? Investigations of how different linguistic interfaces are represented in children's sentence processing system are crucial for our understanding of the human sentence processing system. To explore this question, the present study investigates another interface – the interface between prosody and semantics/pragmatics. More specifically, the present study examines the role of prosody in children's on-line processing of speech act ambiguities by taking advantage of special properties of Mandarin Chinese. We were interested to see whether or not children can use prosodic cues as effectively as adults in resolving speech act ambiguities. To the best of our knowledge, this is the first study investigating children's pragmatic use of prosody in on-line sentence processing where prosody is used to distinguish between two basic communicative acts (i.e., asking a question versus making a statement), as we will discuss below.

The role of prosody in resolving speech act ambiguities

Prosody has several functions that can facilitate utterance comprehension (Cutler, Dahan, & van Donselaar, 1997; Wells, Peppé, & Goulandris, 2004). In the present study, we focused on the pragmatic function of prosody, where prosody helps the hearer understand the speaker's intention (e.g., asking a question versus making a statement) (Verschueren, 1999). The hearer can only base his/her interpretation of the speaker's intention on prosody to determine whether the utterance was a question or a statement.

Mandarin Chinese is ideally suited for evaluating this pragmatic use of prosody in sentence processing, since minimal sentence/prosody pairs are readily available in Mandarin, but rare at best in other languages, including English.[1] The present study takes advantage of the special

[1] In English prosody can also be used to distinguish a question from a statement. For example, English speakers can ask a yes–no question by applying the question intonation to a statement: *John bought a chair?* However, the difference between English and Mandarin is that an English sentence like *John bought a chair* by itself is not ambiguous (i.e., it is a statement when it is in written form) but changing the prosody can make it a question. By contrast, a Mandarin sentence like (3) by itself is ambiguous (i.e., the question and statement interpretations are equally accessible when the sentence is presented visually), and prosody can be used reliably to distinguish between the two interpretations. In daily conversation, Mandarin speakers frequently use sentences like (3) either to ask a *wh*-question or to make a statement, by employing different intonation patterns. Although no studies have specifically looked at how often Mandarin speakers use sentences like (3) to express the two speech acts by applying different intonation patterns, in an interview conducted by the authors where 15 adult Mandarin-speakers were asked whether they would use sentences like (3) to ask a question and whether they would use the same sentences to make a statement, all of them said "yes" to both. The 15 adults were further asked to use sentence (3) either to pose a question or to make a statement. The 8 adults who were asked to pose a question consistently produced the sentence with rising intonation and the 7 adults who were asked to make a statement consistently used level intonation.

properties of Mandarin, where the same sequence of words can be used to perform different speech acts. More specifically, in Mandarin a *wh*-word can be interpreted as a question-marker, or as part of an indefinite noun phrase, which is a constituent of a sentence that makes a statement. Which it is depends on intonation. Consider sentence (3), for example.

(3)	Yuehan	meiyou	chi	shenme	shuiguo[a]
	John	not	eat	what	fruit

a. What fruit did John not eat? (rising intonation on the *wh*-phrase)
b. John didn't eat any fruit. (level intonation on the *wh*-phrase)

[a] One reviewer wanted to know whether test sentences like (3) could also reflect a syntactic ambiguity. As far as we understand, the ambiguity investigated in the present study is generally considered not as a syntactic ambiguity (see e.g., Cheng, 1994), since unlike the ambiguous sentences in previous studies which involve the associations between different elements of the sentences, the test sentences used in this study do not involve such syntactic associations.

Negative sentences with a *wh*-word like (3) are ambiguous in Mandarin Chinese.[2]

This sentence can be used to pose a question, as in (3a): 'What fruit did John not eat?' Alternatively, the same sequence of words can make a statement, as in (3b): 'John didn't eat any fruit'. Prosodic cues are used to distinguish between these two speech acts. A rising intonation on the *wh*-phrase *shenme shuiguo* 'what fruit' indicates the question reading, whereas a level intonation (the absence of rising intonation) on the same *wh*-phrase signals the statement reading. Both kinds of speech acts are used frequently in daily conversation. The intonation contours for the *wh*-phrase *shenme shuiguo* 'what fruit' with rising intonation and with level intonation are illustrated in Fig. 1.

Note that in Mandarin Chinese *wh*-words stay in situ even when they are interpreted as question words. This is probably one of the most important typological features of Mandarin *wh*-questions, that is, unlike many other languages (e.g., English) which form their *wh*-questions by moving a *wh*-phrase to a clause-initial position, Mandarin *wh*-questions are formed by leaving a *wh*-phrase in situ. So Mandarin Chinese is known as a *wh*-in situ language (Huang, 1982a, 1982b; Huang, Li, & Li, 2009). This feature of Mandarin makes the role of prosody especially important in distinguishing a question from a statement, because *wh*-scope is not signalled at the level of syntax (on the surface structure), but at the level of phonology. In addition, prosody in this case serves a cue that carries the illocutionary meaning of an utterance and thus enables the hearer to determine the speaker's communicative intention (i.e., asking a question versus making a statement). We expect that young Mandarin-speaking children might be able to use this prosodic cue as effectively as adults in resolving such ambiguities, since the cue is not just critical for interpretation in Mandarin, but also for distinguishing between two basic speech acts. If this is the case, we might find a prosodic effect that is as robust in children as it is in adults.

To test this prediction, the current experiment examined how Mandarin-speaking adults and 4–5-year-old Mandarin-speaking children interpret sentences like (3) in different prosodic conditions, using the visual world eye-tracking paradigm.

The present study

Method

Participants

Thirty-four monolingual Mandarin-speaking children (mean age 4;8, range 4;1–5;5, 16 girls and 18 boys) and 30 Mandarin-speaking adults (mean age 25, range 23–26, 16 women and 14 men) participated in this experiment. The child participants were recruited from the kindergarten at Beijing Language and Culture University. They had no reported history of speech, hearing or language disorders. The adult participants were students at Beijing Language and Culture University. They had no self-reported speech or hearing disorders. Two adults were excluded from the analysis because we were unable to calibrate them on the eye-tracker. Four children were not included the analysis, because they did not respond correctly to the fill items.

Materials and design

Adult version

A total of 32 target items were constructed each consisting of a picture and two spoken sentences (i.e., the same sentence but with a different intonation pattern: one with rising intonation and one with level intonation). The picture stimuli were always about two characters (either the boy character Xiaoming or the girl character Xiaohong), who are familiar to most children of this age. Xiaoming is a stereotypical boy's name and Xiaohong is a stereotypical girl's name in Mandarin Chinese. The two characters also had stereotypical boy and girl appearances. In each picture, there were always five objects with three objects belonging to one category and two belonging to another category, and the character always chose one object from the category which consists of three objects. In order to control for potential preferences for looking at particular displayed objects, the gender and the position of the character were counterbalanced across trials. The boy character appeared on half of the trials and the girl character on the other half. In addition, on half of the trials the character appeared on the left of the picture and on the other half the character appeared on the right of the picture. Each target sentence contained a subject noun phrase (subject NP), a negative verb phrase (negative VP), a *wh*-word and an object noun phrase (object NP). All the verbs used in the target sentences were action verbs. To describe one picture in detail (see Fig. 2): it showed that there were three types of fruits (bananas, a pear and an orange) and two animals (a

[2] Negation is one of the licensing conditions of the indefinite reading of *wh*-words in Mandarin Chinese. So when *wh*-words like *shei* 'who' and *shenme* 'what' occur in negative sentences, they can be interpreted either as a question-marker, or as an indefinite (roughly like English *some*). But this does not apply to *wh*-words like *na* 'which'. When *na* 'which' appears in negative sentences, it can only be interpreted as a question-marker.

monkey and a cow), and the boy character Xiaoming chose one of the fruits (bananas). For this picture, two sentences were recorded as in (4), one in which the *wh*-phrase *shenme shuiguo* 'what fruit' was produced with rising intonation (Question Prosody hereafter) and the other in which the same *wh*-phrase *shenme shuiguo* 'what fruit' was produced with level intonation (Statement Prosody hereafter). This target sentence consisted of the subject NP *Xiaoming* 'Xiaoming', the negative VP *meiyou zhai* 'not pick', the *wh*-word *shenme* 'what' and the object NP *shuiguo* 'fruit'.

(4) Xiaoming meiyou zhai shenme shuiguo
Xiaoming not pick what fruit
a. What fruit did Xiaoming not pick? (rising intonation on the *wh*-phrase)
b. Xiaoming didn't pick any fruit. (level intonation on the *wh*-phrase)

The 32 target items were divided into two lists with each participant seeing each picture but hearing only one of the two versions of the spoken sentences that could accompany that picture. Target sentences with Question Prosody and those with Statement Prosody were counterbalanced across the two lists with 16 questions and 16 statements in each list. On eight of the statement trials the spoken sentence was a true description of the corresponding picture, and on the other 8 the spoken sentence was a false description of the corresponding picture.

In addition, 32 filler items were added to each experimental list: 16 questions and 16 statements. Each filler item consists of a picture and a spoken sentence. The picture stimuli were similar to those used in the target items (An example can be found in Appendix A). The spoken sentence in each item was either an unambiguous question or a statement. Two examples are given in (5) and (6).

(5) Xiaohong meiyou mo naxie dongwu?[a]
Xiaohong not touch which animal
'Which animals did Xiaohong not touch?'

[a] Sentence (5) is always a question regardless of what intonation is placed on the wh-phrase naxie dongwu 'which animals'.

(6) Xiaohong meiyou chi renhe dongxi.
Xiaohong not eat any thing
'Xiaohong didn't eat any food.'

On half of the statement trials the spoken sentence was a true description of the corresponding picture, and on the other half the spoken sentence was a false description of the picture.

In each experimental list, the 32 target and 32 filler items were arranged in random order. Adult participants were randomly assigned to one of the two lists, with 14 participants run on each list.

Child version

Eighteen out of the 32 target items and 12 out of the 32 filler items in the adult version were used for children. So the child experimental session consists of 30 items: nine target sentences with Question Prosody, nine target sentences with Statement Prosody, six filler questions and six filler statements. Four target and three filler statements are true descriptions of the corresponding pictures and the other five target and three filler statements are false descriptions of the corresponding pictures. In the adult

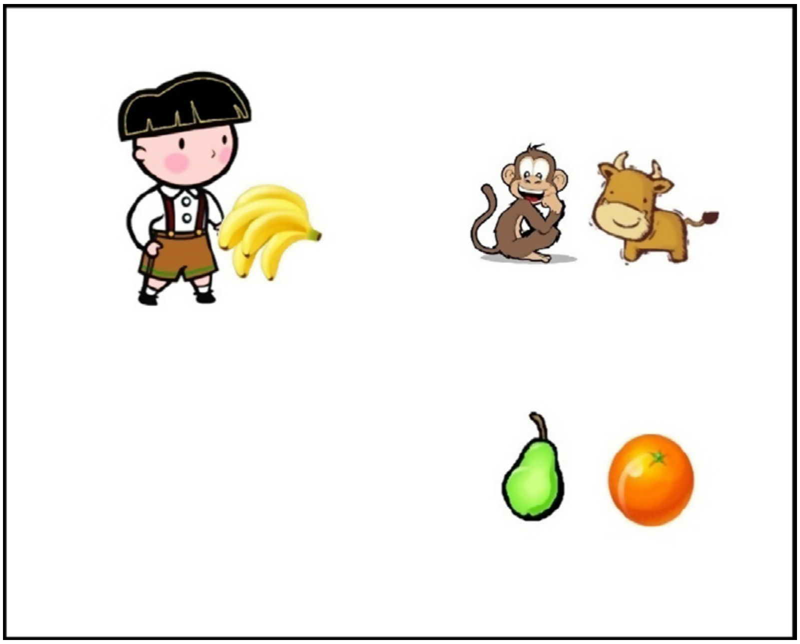


Fig. 2. Example target picture used in the experiment.

version, prosody was manipulated within participants. In this child version, prosody was also manipulated within participants but was blocked. This manoeuvre was taken to maximise the chances of detecting the prosodic effect in children. Previous research found that when sentences with different prosodic versions were presented to children in the same experimental session, a contamination effect might occur, that is, a strong tendency to perseverate across trials could easily wipe out a small or fragile effect of prosody (see Snedeker & Yuan, 2008). In order to avoid this possible contamination, we used a blocked deign. Two counterbalanced experimental lists were constructed. Each list consists of two experimental blocks. The first block of one list contained target sentences with Question Prosody followed by the second block containing sentences with Statement Prosody, while the first block of the other list contained target sentences with Statement Prosody followed by the second block consisting of sentences with Question Prosody. In each block, the target trials were interspersed with filler trials. Target trials containing sentences with Question Prosody were interspersed with filler trials containing filler statements, whereas target trials containing sentences with Statement Prosody were interspersed with filler trials containing filler questions. Child participants were randomly assigned to one of the two lists, with 17 participants run on each list. Thus, half of the children heard the target questions first followed by target statements. The other half of the children heard the two versions of the sentences in reverse order. A full list of the target sentences in the child version is included in Appendix B.

The production of the test stimuli

The test sentences were produced by a female native speaker of Beijing Mandarin. She was asked to produce the test sentences in a child-directed manner. The recording was conducted in a sound-attenuated recording booth at Macquarie University. Two versions of each target sentence were recorded, one with rising intonation on the *wh*-phrase and one with level intonation on the same *wh*-phrase. To determine whether the target sentences were produced successfully, we did an acoustic analysis of each element in the target sentences. Each sentence contains a subject NP, a negative VP, a *wh*-word and an object NP. The duration and intensity of each element was measured and paired *t*-tests were conducted to verify the differences between the two versions of target sentences (see Table 1).

As indicated in Table 1, no significant differences were found in the elements before the *wh*-word. A significant difference was observed in the *wh*-words between the two prosodic conditions. The *wh*-words with Question Prosody were longer than their counterparts with Statement Prosody ($p < .001$). The *wh*-words with Statement Prosody had greater intensity than their counterparts with Question Prosody ($p = .01$). The results of the acoustic analysis confirmed that the target sentences were produced successfully. The difference between the two prosodic conditions lies in the different acoustic features of the *wh*-words.

Table 1
(a) Duration analyses of the stimuli in the experiment (standard deviation in parentheses). (b) Intensity analyses of the stimuli in the experiment (standard deviation in parentheses).

Dependent variable	Mean for Question Prosody	Mean for Statement Prosody	Analysis
(a)			
Subject NP	820 ms (28)	817 ms (27)	$t(31) = .40$, $p = .70$
Pause	501 ms (26)	501 ms (20)	$t(31) = .03$, $p = .98$
Negative VP	909 ms (21)	906 ms (23)	$t(31) = 1.14$, $p = .26$
Pause	485 ms (25)	510 ms (26)	$t(31) = 1.56$, $p = .13$
Wh-word	774 ms (23)	665 ms (14)	$t(31) = 12.32$, $p < .001^{**}$
Pause	483 ms (24)	502 ms (27)	$t(31) = 1.82$, $p = .08$
Object NP	927 ms (27)	936 ms (22)	$t(31) = .59$, $p = .56$
(b)			
Subject NP	65.95 dB (2)	65.48 dB (2)	$t(31) = .82$, $p = .42$
Negative VP	66.23 dB (1)	66.29 dB (1)	$t(31) = .18$, $p = .86$
Wh-word	65.58 dB (1)	66.93 dB (2)	$t(31) = 2.67$, $p = .01^{*}$
Object NP	64.94 dB (1)	64.93 dB (1)	$t(31) = .01$, $p = .99$

Procedure

Both children and adults were tested using the visual world paradigm (Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995). This paradigm is based on an early observation by Cooper (1974), who pointed out that when participants are simultaneously presented with spoken language while viewing a visual scene, their eye movements are very closely synchronised to the referential processing of the concurrent linguistic input. This linguistic sensitivity of this paradigm has been validated in various studies by Tanenhaus and colleagues (e.g., Allopenna, Magnuson, & Tanenhaus, 1998; Sedivy, Tanenhaus, Chambers, & Carlson, 1999). This paradigm has also been successfully used to test children's linguistic knowledge in on-line sentence processing (e.g., Arnold, Brown-Schmidt, & Trueswell, 2007; Choi & Trueswell, 2010; Höhle et al., 2009; Sekerina & Trueswell, in press; Snedeker & Trueswell, 2004; Snedeker & Yuan, 2008; Trueswell, Sekerina, Hill, & Logrip, 1999; Zhou et al., in press). In this study, participants were presented with a spoken sentence while viewing a picture. Participants were told that they were going to see some pictures, and after each picture, the puppet, Kermit the Frog, was going to explain to them what he thought had happened in the picture. It was made clear to the participant that the puppet did not always pay close attention to the picture and thus was sometimes unsure about what happened in the picture. If that was the case, the puppet would make a guess about what happened in the picture or ask the participant a question. On each trial, the participant's task was to decide whether the puppet accurately said what happened in the picture or asked a question about the picture. Whenever the puppet said what happened in the picture, the participant was instructed to judge whether the puppet was right or wrong. But if the puppet asked a question about the picture, the participant was instructed to answer the question. This Question-Statement task has been successfully used with Mandarin-speaking children (Zhou & Crain, 2011). Participants' eye movements were recorded using an EyeLink 1000 eye tracker (by SR Research Ltd., Mississauga, Ontario, Canada) interfaced with a PC computer. The EyeLink 1000 allows remote eye tracking, without a head support. The eyetracker provides information about the participant's point of gaze at a sampling rate of 500 Hz. The eyetracker has accuracy of 0.5° of visual angle. The picture stimuli were displayed on the monitor. Spoken test sentences were presented to the participants through the PC computer connected to two external speakers. Though the eye tracker does not require head stabilization, the child participants were held still by an adult experimenter, and they leaned slightly back in a chair in front of the monitor. This manoeuvre was taken to reduce back and forth movements by the child participants. The distance between the participants' eyes and the monitor was about 60 cm.

Before the actual experiment, we had an introduction session to familiarise the child participants with the experimental procedure as well as the objects that were presented in the pictures. The introduction session was followed by the experimental session. In the adult version of the experiment, the experimental session began with four practice trials followed by 64 test trials (32 target and 32 filler trials). In the child version of the experiment, the experimental session consists of four practice trials and 30 test trials in two blocks (18 target and 12 filler trials). The spoken sentences on the practice trials were composed of two questions and two statements. Before each trial, a picture of Kermit the Frog was presented at the centre of the monitor, which anchored the beginning of each trial, and served to capture the participants' attention. The picture was also used to make it appear that the puppet was talking. This picture gave way to the trial as soon as the participant focused on the centre of the monitor.

The spoken sentence started 2000 ms after the appearance of the picture stimulus. Participants' eye movements were recorded for 6 s from the onset of the subject NP. The mean length of the target sentences was 4868 ms for adults, and 4885 ms for children. The mean length of each element of the target sentences in the two prosodic conditions was illustrated in Fig. 3 (child version). See Table 1 for the adult version.

Data treatment

We only included a participant's data in the analysis if he/she performed above chance on the filler trials, where the spoken sentences were unambiguous questions or statements. Four children were excluded from the analysis because they did not respond correctly to the fill items. Two adults were excluded because we were unable to calibrate them on the eyetracker. The remaining 28 adults and 30 children were include in the analysis. Trials were

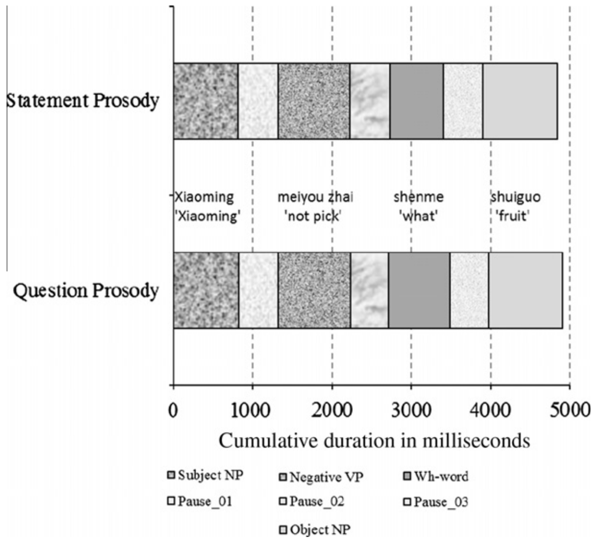


Fig. 3. Time courses for the target sentences in the child version.

excluded if there was more than 33% track loss from the onset of the *wh*-word until the sentence was completed. Three trials were excluded for adults and no trials were excluded for children.

There were two sets of data: the off-line responses and the eye-movement data. The off-line responses consist of participants' responses to each experimental trial. In analysing the eye-movement data, participants' fixations were coded in three categories: the statement-compatible area (objects corresponding to the statement interpretation), the question-compatible area (objects corresponding to the question interpretation) and the irrelevant area (objects corresponding to neither interpretation). Use example (4) (repeated here as (7)) and Fig. 4 to illustrate.

(7) Xiaoming meiyou zhai shenme shuiguo
Xiaoming not pick what fruit
What fruit did Xiaoming not pick? (rising intonation on the *wh*-phrase)
Xiaoming didn't pick any fruit. (level intonation on the *wh*-phrase)

In this example, the statement-compatible area is I (i.e., Xiaoming and the bananas), the question-compatible area is III (i.e., the pear and the orange), and the irrelevant area is II (i.e., the monkey and the cow). The basic idea is that if sentence (7) is interpreted as a statement, meaning: 'John didn't pick any fruit', then participants' judgement of this statement should be based on their observation in area I (e.g., a rejection like "It's wrong, because Xiaoming picked bananas"). If, on the other hand, the sentence is interpreted as a question, asking: 'What fruit did Xiaoming not pick?', participants' answers to this question should be based on their observation in area III (e.g., an answer like "a pear and an orange"). The proportion of fixations following the onset of the subject NP for each category was computed in a time window of 6000 ms. The time window was made about 1000 ms longer than the mean length of the spoken

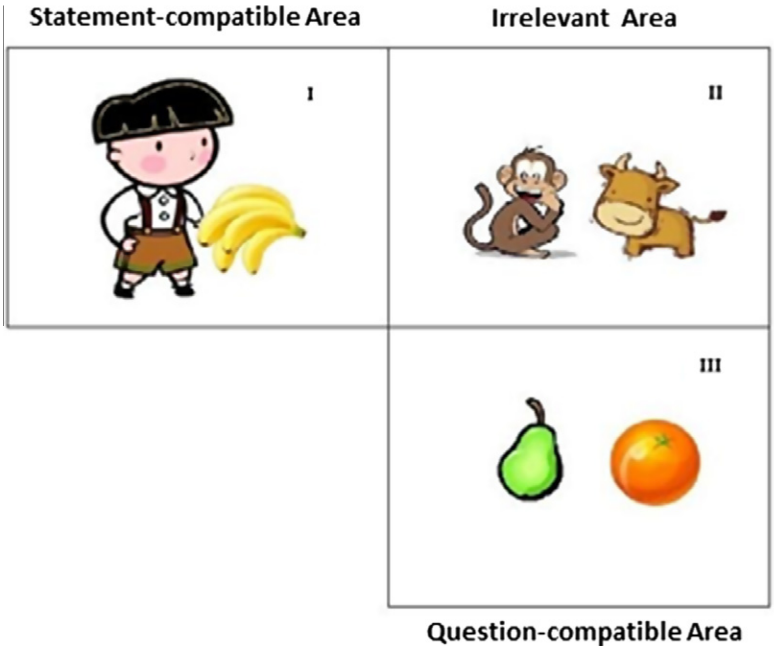


Fig. 4. Example of interest areas.

sentences, so that participants would have sufficient time to process and integrate different levels of linguistic representations when comprehending the spoken sentences.

Predictions

With respect to the off-line responses, if children were sensitive to the prosodic cues (rising intonation versus level intonation), and could use the cues to disambiguate between the two speech acts (asking a question versus making a statement), they would be expected to provide an answer to a target sentence with Question Prosody, since with rising intonation on the *wh*-phrase the sentence indicated a question. By contrast, children should judge a target sentence with Statement Prosody to be either true or false, since with level intonation on the *wh*-phrase the sentence expressed a statement. Consider sentence (7) and Fig. 4, for example. Children should respond to the sentence with Question Prosody by offering an answer "a pear and an orange", and they should respond to the sentence with Statement Prosody by rejecting it on the grounds that Xiaoming picked bananas.

Concerning the eye gaze data, if children were sensitive to the intonational cues and were able to use the cues in on-line sentence processing, we would expect them to look more to the question-compatible area when they listened to target sentences with Question Prosody than when listening to sentences with Statement Prosody. By contrast, they would be expected to fixate more on the statement-compatible area when they were presented with target sentences with Statement Prosody than when presented with sentences with Question Prosody. This expected difference between the two conditions should occur after they heard the *wh*-word, since the prosodic features of the two conditions started to differ from the onset of the *wh*-word, as indicated in Table 1. On the example trial, more fixations would be expected in area III (i.e., the pear and the orange) in the Question Prosody condition than in the Statement Prosody condition, whereas an opposite pattern should be observed in area I (i.e., Xiaoming and the bananas).

Results and discussion

Eye-movement data

The proportion of fixations following the onset of the subject NP was computed in a time window of 6000 ms for the two critical categories: the statement-compatible area (I) and the question-compatible area (III). This 6000 ms time window was broken down into 20 segments, each with a duration of 300 ms. In the adult version, the *wh*-word started 2735 ms after the onset of the subject NP and finished at 3399 ms. In the child version, the *wh*-word started 2746 ms after the onset of the subject NP and finished at 3416 ms. To provide an overview of the eye-movement data, the results are first presented in the form of descriptive graphs followed by more detailed statistical analyses.

Fig. 5 shows the proportion of fixations of adults (upper panel) and children (lower panel) in the question-compatible area (III) across the two prosodic conditions. Fig. 6 summarises the proportion of fixations of adults and children in the statement-compatible area (I) across the two prosodic conditions. The two figures indicate that adults and children exhibited similar eye gaze patterns in the two critical areas in the two prosodic conditions. For both adults and children, a higher proportion of fixations was found in the question compatible area (e.g., the pear and the orange) for sentences with Question Prosody (i.e.,

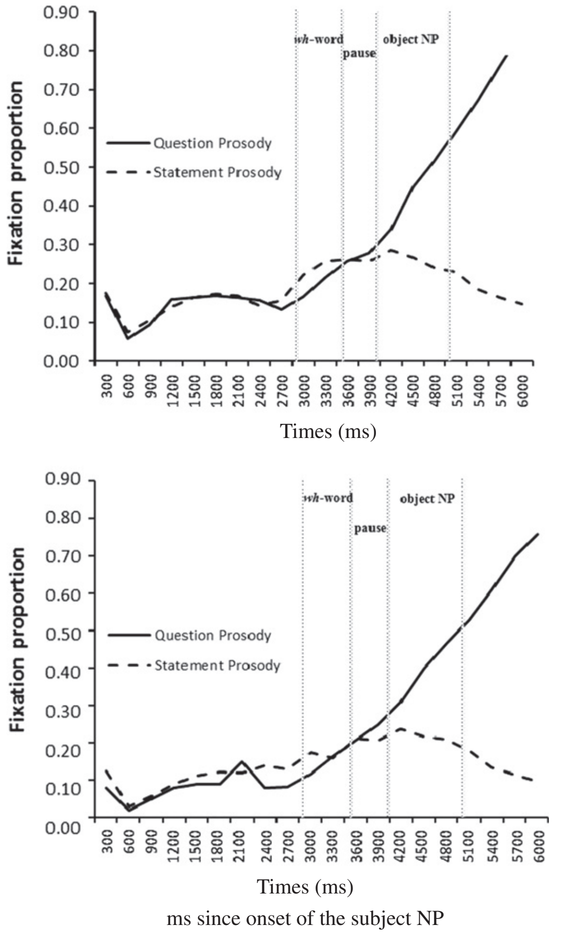


Fig. 5. Average fixation proportions over time in the question-compatible area (III) in the two prosodic conditions, adults (upper panel) and children (lower panel).

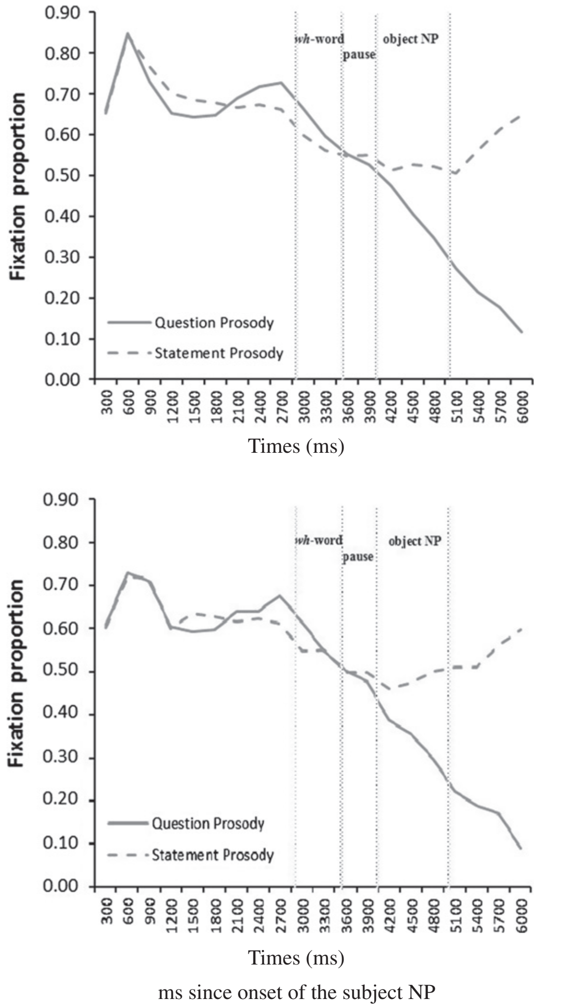


Fig. 6. Average fixation proportions over time in the statement-compatible area (I) in the two prosodic conditions, adults (upper panel) and children (lower panel).

rising intonation on the *wh*-phrase) than for sentences with Statement Prosody (i.e., level intonation on the *wh*-phrase). The adults showed increased looks to the question-compatible area from the onset of the pause after the *wh*-word (in the time span between 3600 and 3900 ms) and this pattern became increasingly clear on hearing the object NP after the pause (with increasingly more fixations for sentences with Question Prosody and less fixations for sentences with Statement Prosody). The children exhibited similar eye-movement patterns (see Fig. 5). By contrast, for both adults and children an opposite pattern was observed in the statement-compatible area (e.g., Xiaoming and the bananas). There is a higher probability of looks in the statement-compatible area for sentences with Statement Prosody than for sentences with Question Prosody. Both children and adults started to look more to the statement-compatible area from the onset of the pause after the *wh*-word (in the time span between 3600 and 3900 ms), and on encountering the object NP a sharp decrease of looks was observed for sentences with Question Prosody and a steady increase of looks was observed for sentences with Statement Prosody (see Fig. 6).

To assess these fixation patterns statistically, mixed-effects logistic regression models were applied using the R software package, version 2.13.2 (R Development Core Team, 2011). We fit the data for adults and children separately, focusing on the looks to the two critical categories (the question-compatible area and the statement-compatible area) between 3600 and 6000 ms (eight time segments of 300 ms from the onset of the pause after the *wh*-word). The models treated prosodic condition (Question Prosody versus Statement Prosody) and time as fixed effects, with random intercepts and slopes for both participants and items (Baayen, Davidson, & Bates, 2008). Because fixation proportions were used, the data were first transformed using an empirical logit function (Barr, 2008; Jaeger, 2008).[3]

[3] In the logistic terms, the data expresses the changes in fixation likelihood of looking at versus not looking at an given interest area. According to Jaeger (2008), empirical logit transformation can correct the problem of heterogeneity of variance distribution in using proportional data.

Table 2
Fixed effects from best-fitting mixed-effects logistic regression model of probability of looks to the question-compatible area and the statement-compatible area (empirical logit tramsformed), adults.

Category	Fixed effects	Estimate	SE	*t* Value
Question-compatible area				
	(Intercept)	−.25	.09	−2.52*
	Condition (Statement Prosody)	−.45	.10	−4.29***
	Time	.22	.01	16.01***
	Condition (Statement Prosody) × Time	−.26	.02	−15.11***
Statement-compatible area				
	(Intercept)	−.22	.09	−2.50*
	Condition (Statement Prosody)	.31	.09	3.62***
	Time	−.16	.02	−10.60***
	Condition (Statement Prosody) × Time	.19	.02	11.91***
Formula in *R*: fixation ∼ condition ∗ time + (1 + condition + time|participant) + (1 + condition + time|item)				

* $p < .05$.
*** $p < .001$.

Table 3
Fixed effects from best-fitting mixed-effects logistic regression model of probability of looks to the question-compatible area and the statement-compatible area (empirical logit tramsformed), children.

Category	Fixed effects	Estimate	SE	*t* Value
Question-compatible area				
	(Intercept)	−.21	.10	−2.13*
	Condition (Statement Prosody)	−.55	.12	−4.71***
	Time	.23	.02	15.22***
	Condition (Statement Prosody) × Time	−.27	.02	−13.59***
Statement-compatible area				
	(Intercept)	−.26	.08	−3.14**
	Condition (Statement Prosody)	.38	.10	3.80***
	Time	−.17	.01	−11.75***
	Condition (Statement Prosody) × Time	.19	.02	12.76***
Formula in *R*: fixation ∼ condition ∗ time + (1 + condition + time|participant) + (1 + condition + time|item)				

* $p < .05$.
** $p < .01$.
*** $p < .001$.

Table 2 summarises the model results for the adult data in the two critical categories (the question-compatible area and the statement-compatible area). In the question-compatible area, the negative coefficient for the significant main effect of prosodic condition (='Condition (Statement Prosody)') reflects the fact that the adults looked significantly more at the question-compatible area (between 3600 and 6000 ms from the pause after the *wh*-word) in the Question Prosody condition than in the Statement Prosody condition. Time is also a reliable predictor and it interacts with prosodic condition (='Condition (Statement Prosody) × Time') such that the probability of looking at the question-compatible area decreased over time for sentences with Statement Prosody from the onset of the pause after the *wh*-word. By contrast, an opposite pattern was found in the statement-compatible area. Prosodic condition is a reliable positive predictor in this case, indicating that the adults fixated more often on the statement-compatible area in the Statement Prosody condition than in the Question Prosody condition over the same time period. The positive coefficient for the interaction between time and prosodic condition confirms that the probability of fixating on the statement-compatible area increased over time for sentences with Statement Prosody.

Children exhibited similar patterns. The fixed effects from the models fitted to the child data appear in Table 3. In the question-compatible area, children, like adults, exhibited a main effect of prosodic condition. The negative coefficient for this effect indicates that they looked significantly more at the question-compatible area when listening to sentences with Question Prosody than when listening to sentences with Statement Prosody (in the time period between 3600 and 6000 ms from the onset of the pause after the *wh*-word). In addition, a main effect of time and an interaction between time and prosodic condition were also found in children. The negative coefficient for the interaction confirms that the probability of looking at the question-compatible area decreased over time for sentences with Statement Prosody. This pattern was reversed in the statement-compatible area, however. The children showed increased looks to the statement-compatible area for sentences with Statement Prosody than for sentences with Question Prosody over the same time period, as indicated by the positive coefficient for the main effect of prosodic

Table 4
Fixed effects from best-fitting mixed-effects logistic regression model of probability of looks to the question-compatible area and the statement-compatible area (empirical logit tramsformed), adult and child data combined.

Category	Fixed effects	Estimate	SE	*t* Value
Question-compatible area				
	(Intercept)	−.26	.09	−2.76**
	Condition (Statement Prosody)	−.44	.11	−4.11***
	Time	.22	.01	16.34***
	Group (Children)	.04	.06	0.60
	Condition (Statement Prosody) × Time	−.26	.02	−15.89***
	Condition (Statement Prosody) × Group (Children)	−.12	.09	−1.37
	Time × Group (Children)	.02	.01	1.87
	Condition (Statement Prosody) × Time × Group (Children)	−.02	.02	−1.43
Statement-compatible area				
	(Intercept)	−.22	.09	−2.49*
	Condition (Statement Prosody)	.30	.09	3.27**
	Time	−.16	.01	−11.34***
	Group (Children)	−.04	.06	−0.71
	Condition (Statement Prosody) × Time	.19	.01	13.54***
	Condition (Statement Prosody) × Group (Children)	.11	.09	1.20
	Time × Group (Children)	−.01	.01	−1.09
	Condition (Statement Prosody) × Time × Group (Children)	.01	.02	0.27

Formula in *R*: fixation ∼ condition ∗ time ∗ group + (1 + condition + time|participant) + (1 + condition + group + time|item)

* $p < .05$.
** $p < .01$.
*** $p < .001$.

condition. The positive coefficient for the significant interaction between time and prosodic condition confirms that the probability of fixating on the statement-compatible area increased over time for sentences with Statement Prosody from the onset of the pause after the *wh*-word.

The patterns displayed by Figs. 5 and 6 were supported by the statistical modelling. Adults and children showed similar eye-movement patterns in the two critical categories in the two prosodic conditions. Prosodic condition and time are reliable predictors of the fixation patterns of both adults and children. In order to see whether there exist any differences between adults and children, new models were developed in which group (adults versus children) was included as an experimental factor. More specifically, models were fitted to the entire data set (both adult and child data) for the two critical categories, treating prosodic condition (Question Prosody versus Statement Prosody), time and group (adults versus children) as fixed effects, with random intercepts and slopes for both participants and items. The results are summarised in Table 4. In both the question-compatible and statement-compatible areas, no significant effects of group (='Group (Children)') were observed, which confirms that children and adults exhibited similar fixation patterns in both areas. The negative coefficient for the main effect of prosodic condition in the question-compatible area indicates that both groups exhibited a higher probability of looks to the question-compatible area for sentences with Question Prosody than for sentences with Statement Prosody. The probability of fixating on the question-compatible area decreased over time for sentences with Statement Prosody, as indicated by the negative coefficient for the significant interaction between time and prosodic condition. Both groups showed an opposite fixation pattern in the statement-compatible area. The positive coefficient for the main effect of prosodic condition reflects that for both adults and children there were increased looks to the statement-compatible area in the Statement Prosody condition than in the Question Prosody condition. In addition, the positive coefficient for the interaction between time and prosodic condition confirms that for both groups the probability of fixating on the statement-compatible area increased over time for sentences with Statement Prosody.

To summarise, the eye-movement data show that both adults and children launched more fixations to the question-compatible area after hearing the *wh*-word with rising intonation than after hearing the *wh*-word with level intonation, whereas both groups looked more to the statement-compatible area after hearing the *wh*-word with level intonation than after hearing the *wh*-word with rising intonation. The prosodic effects found in children were as robust as those in adults, as illustrated by the statistical models discussed above. Both adults and children exhibited the expected eye movement patterns based on the prosodic cues provided, and the expected patterns occurred even before the onset of the object NP (e.g., *shuiguo* 'fruit') of the *wh*-phrase (e.g., *shenme shuiguo* 'what fruit'). The findings suggest that not only were the children sensitive to the prosodic cues (rising intonation versus level intonation), but they were able to use the cues as effectively as adults in the on-line processing of speech acts (asking a question versus making a statement).

In order to see whether children could incorporate this on-line sensitivity in their off-line judgements, we analysed the off-line responses of the participants.[4]

[4] The discrepancy between off-line and on-line tasks has been reported in several previous studies (e.g., Brandt-Kobele & Höhle, 2010; Höhle et al., 2009; Sekerina, Stromswold, & Hestvik, 2004; Zhou et al., in press). These studies found that on-line tasks like eye-tracking can reveal a more adult-like linguistic competence than off-line tasks.

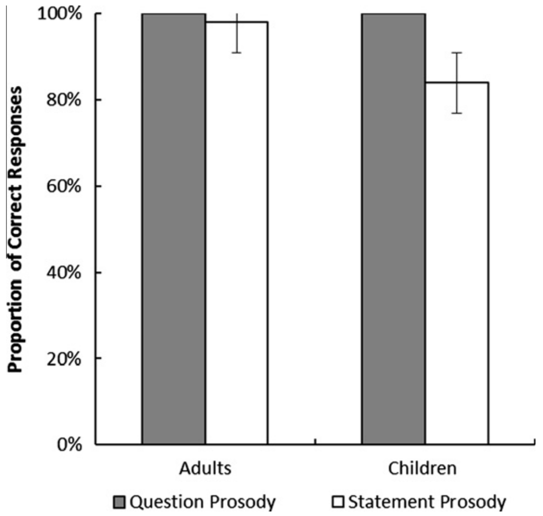


Fig. 7. Mean proportion of correct responses by adults and children in the two prosodic conditions.

Off-line responses

We recorded the responses of the participants in the two prosodic conditions. Analyses were conducted on the proportion of correct responses in which the judgments accurately reflected the prosody of the spoken sentence. Consider sentence (7) and Fig. 4, for example. A correct response to sentence (7) with Question Prosody is a response by offering an answer like "a pear and an orange" (question-compatible response hereafter), and a correct response to the sentence with Statement Prosody is a response by rejecting the sentence on the grounds that Xiaoming picked bananas (statement-compatible response hereafter). Fig. 7 gives the proportion of correct responses by adults and children in the two prosodic conditions.

In response to target sentences with Question Prosody, both adults and children provided question-compatible responses 100% of the time. For example, both adults and children responded to sentence (7) with rising intonation on the *wh*-phrase *shenme shuiguo* 'what fruit' by providing an answer "li he juzi" 'a pear and an orange'.

In response to target sentences with Statement Prosody, the adults provided correct statement-compatible responses 98% of the time and the children 84% of the time. For example, when rejecting sentence (7) with level intonation on the *wh*-phrase *shenme shuiguo* 'what fruit', both adults and children pointed out that Kermit was wrong by saying "budui, Xiaoming zhai-le xiangjiao" 'wrong, Xiaoming picked bananas'. A Mann–Whitney test showed that the adults gave correct statement-compatible responses to a much higher degree than the children did ($z = 4.26$, $p < .001$). A closer inspection of the responses found that the difference between adults and children was mainly in their responses to the false target sentences. In response to true target sentences, both adults and children accepted the sentences 100% of the time (adults: 224/224 trials; children: 120/120 trials). In response to false target sentences, the adults correctly rejected the sentences 96.43% of the time (216/224 trials)[5] and the children did so 71.33% of the time (107/150 trials). The other 28.67% of the time (43/150 trials) the children incorrectly judged the false target sentences to be true, e.g., they responded to sentence (7) with Statement Prosody by saying "dui" 'right'. We will return to this difference between children and adults in the general discussion. Nevertheless, none of the child participants gave question-compatible responses in the Statement Prosody condition. The results demonstrate that children, like adults, were relying on the prosodic cues (rising intonation versus level intonation) when making judgements about whether a given sentence was a question or a statement. Children showed sensitivity to the prosodic cues in distinguishing between a question and a statement, both on-line and off-line.

General discussion

The present study sought to investigate whether young Mandarin-speaking children are sensitive to prosodic cues in on-line processing of speech act ambiguities. To assess this, we examined the fixation patterns of both children and adults in the critical areas after hearing the *wh*-word with different prosodic features (rising intonation versus level intonation). We found that 4-year-old Mandarin-speaking children exhibited similar fixation patterns to adults in the two critical areas. Hearing the *wh*-word with rising intonation triggered increased looks to the question-compatible area and hearing the *wh*-word with level intonation triggered increased looks to the statement-compatible area. This prosodic effect found in children was as robust as it was in adults. In order to see whether children could incorporate this on-line sensitivity in making their off-line judgements, their off-line responses were analysed in the two prosodic conditions. It was found that children's judgments accurately reflected the prosody of the spoken sentences. Rising intonation on the *wh*-word led to question-compatible responses and level intonation on the *wh*-word led to statement-compatible responses. This is evidence that children know that rising intonation on the *wh*-word turns sentences into questions, whereas level intonation on the *wh*-word turns them into statements. Children reliably used prosodic cues to guide their off-line comprehension of speech acts.

Taken together, the results of the present study suggest that young Mandarin-speaking children are sensitive to prosodic information in resolving speech act ambiguities, both on-line and off-line. Mandarin-speaking children can use prosodic information as effectively as adults in their on-line processing of speech acts as indicated by the eye-movement data. This finding stands in contrast to previous studies showing that young children are sensitive to prosodic information in on-line sentence processing, but they use such information less effectively than adults, as evidenced by the fact that children showed delayed effect of prosody, as compared to adults (Ito et al., in press; Sekerina & Trueswell, in press; Snedeker

[5] The other 3.57% of the time (8/224 trials) adults incorrectly accepted the false target sentences.

& Yuan, 2008; Zhou et al., in press). However, the present study, for the first time, found a prosodic effect that was as robust in children as it was in adults. This finding invites the conclusion that young Mandarin-speaking children are as good as adults in the pragmatic use of prosody in on-line sentence processing. Our finding has two important implications in understanding the role of prosody in children's on-line sentence processing.

First, our findings provide insight into how different linguistic interfaces are represented in children's sentence processing system. Previous studies have shown that children are not as adept as adults at using prosodic cues to resolve syntactic ambiguities, which we attribute to the less automatic mapping between prosody and syntax in child language. Unlike previous studies, the present study, for the first time, found that children use prosodic cues as effectively as adults in resolving speech act ambiguities. This finding suggests that the mapping between prosody and semantics/pragmatics in child language is as well-established as in adult language. On the basis of these findings, it seems safe to say that linguistic interfaces are not equally represented in children's sentence processing system. The interface between prosody and semantics/pragmatics is better established than the interface between prosody and syntax in this system.

However, we also need to consider other factors that might have contributed to the better performance of the children in our study than those in previous studies. For example, our study used different prosodic cues than those in previous studies. In our study, intonational cues (rising intonation versus level intonation) were used as prosodic cues, but previous studies used either prosodic boundary or pitch accent as prosodic cues. So it might be hypothesised that intonational cues are easier for young children than prosodic boundary or pitch accent. Although we do not have direct evidence as to whether children can process intonation more easily than prosodic boundary or pitch accent, previous studies seem to suggest that young children have no difficulty in perceiving prosodic boundary or pitch accent at the word level. The first evidence is from Choi and Mazuka (2003). As discussed in the introduction, Choi and Mazuka found that 3–4-year-old Korean-speaking children were able to use prosodic boundary as cue to resolve word-segmentation ambiguity. Another evidence is from Grassmann and Tomasello (2007, 2010). Grassmann and Tomasello found that children as young as 2 years of age can detect pitch accent in a sentence and understand that pitch accent is used to indicate referential newness in the sentence. Based on these findings, it seems quite unlikely that the different prosodic characteristics used in our study than in previous studies could have led to the better performance of the children in our study.

Second, the findings of the current study led us to propose that we also need to take into account the relevance of the cues to the intended interpretation when considering the role of prosody in sentence comprehension. In other words, we need to consider how important the prosodic cue is in leading to a correct interpretation of the sentence. In the present study, prosodic cues were the sole available cue to the intended interpretation, which means that children had to rely on the prosodic cues in order to arrive at the intended interpretation. In addition, prosody serves as a cue that carries the illocutionary meaning of an utterance, namely, it enables the hearer to determine the speaker's communicative intention, i.e., whether the speaker is asking a question or making a statement. This pragmatic use of prosody should be acquired early in the course of language development, since it is important for the development of general communication skills, e.g., making a request, asking a question, etc. Moreover, as discussed in the introduction, Mandarin Chinese is a *wh*-in situ language where *wh*-questions are formed by leaving a *wh*-phrase in situ. In other words, *wh*-words always stay in situ regardless of whether they occur in questions or statements. This feature of Mandarin makes the role of prosody especially important in distinguishing a question from a statement, because *wh*-scope is not signalled at the level of syntax (on the surface structure), but at the level of phonology. Given the importance of the prosodic cues in understanding the meanings of the target sentences, it is not surprising that young Mandarin-speaking children in our study exhibited adult-like sensitivity to the prosodic cues in sentence processing. The effects of such language-specificity have so far not been explored in the investigation of the role of prosody in children's on-line sentence processing.

At this point, one question remains to be answered: how can we account for the difference between adults and children in their responses to the false target sentences in the Statement Prosody condition? We found that the adults correctly rejected the false target sentences to a much higher degree than the children did, though both the adults and the children interpreted the target sentences as statements in this condition. The adults correctly rejected the false target sentences 96.43% of the time and the children did so 71.33% of the time. The other 28.67% of the time the children incorrectly judged the false target sentences to be true. This difference between adults and children suggest that children are more attempted to say ''yes'' to the false target sentences. They felt more reluctant to switch from saying ''yes'' to saying ''no'' in responding to the false target sentences. We speculate that children's difficulty in switching between a ''yes'' and a ''no'' response might be due to children's limited cognitive control abilities (e.g., Choi & Trueswell, 2010; Clackson, Felser, & Clahsen, 2011). Adults are able to exert rapid control over their responses, for example, to switch quickly between ''yes'' and ''no'' responses. Children, however, may have difficulty in exerting control over their switch from a ''yes'' response to a ''no'' response if they have a strong tendency to say ''yes''.

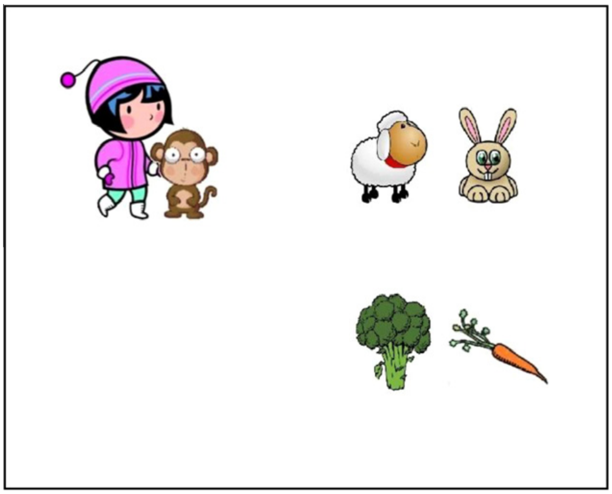
To conclude, our study shows that children are as good as adults in the pragmatic use of prosody in on-line sentence processing. This finding suggests that children have a sentence processing system that can use prosodic information rapidly and effectively to resolve ambiguities involving the pragmatic use of prosody. Of course, we cannot simply generalise our results to other ambiguities involving the interface between prosody and semantics/pragmatics. Further investigations of different types of pragmatic use of prosody are required. The present study is an important beginning of an effort to understand how different linguistic interfaces are represented in children's sentence processing system. In addition, our study only focused on 4–5-year-old children and leaves open the question of whether younger children are also capable of using this prosodic information to resolve the ambiguity. It would be useful to begin to map out the developmental time course of this pragmatic use of prosody in children's sentence processing, by examining younger children using similar tasks. This is a necessary next step in the evaluation of our proposal.

Acknowledgements

This work was supported by the Macquarie University Research Fellowship to the first author, and two Australian Research Council Discovery Grants (DP1096160 & DP0879842). We are grateful to two anonymous reviewers and the editor for their insightful comments and suggestions on an earlier version of the paper. We would also like to thank the children and adults who took part in the experiments for their time and patience.

Appendix A

Example filler picture used in the experiment



Appendix B

Target sentences in the child version (each sentence was recorded in two prosodic versions one in which rising intonation was placed on the *wh*-phrase and the other in which level intonation was placed on the *wh*-phrase)

(1)	Xiaoming	meiyou	zhai	shenme	shuiguo
	Xiaoming	not	pick	what	fruit
(2)	Xiaoming	meiyou	xuan	shenme	dongwu
	Xiaoming	not	choose	what	animal
(3)	Xiaoming	meiyou	wan	shenme	wanju
	Xiaoming	not	play	what	toy
(4)	Xiaoming	meiyou	mai	shenme	chide
	Xiaoming	not	buy	what	food
(5)	Xiaoming	meiyou	mo	shenme	dongwu
	Xiaoming	not	touch	what	animal
(6)	Xiaoming	meiyou	chi	shenme	shuiguo
	Xiaoming	not	eat	what	fruit
(7)	Xiaoming	meiyou	xuan	shenme	yifu
	Xiaoming	not	choose	what	clothes
(8)	Xiaoming	meiyou	chi	shenme	shucai
	Xiaoming	not	eat	what	vegetable
(9)	Xiaoming	meiyou	wei	shenme	dongwu
	Xiaoming	not	feed	what	animal
(10)	Xiaohong	meiyou	zhai	shenme	shuiguo
	Xiaohong	not	pick	what	fruit
(11)	Xiaohong	meiyou	xuan	shenme	dongwu
	Xiaohong	not	choose	what	animal
(12)	Xiaohong	meiyou	wan	shenme	wanju
	Xiaohong	not	play	what	toy
(13)	Xiaohong	meiyou	mai	shenme	chide
	Xiaohong	not	buy	what	food
(14)	Xiaohong	meiyou	mo	shenme	dongwu
	Xiaohong	not	touch	what	animal
(15)	Xiaohong	meiyou	chi	shenme	shuiguo
	Xiaohong	not	eat	what	fruit
(16)	Xiaohong	meiyou	xuan	shenme	yifu
	Xiaohong	not	choose	what	clothes
(17)	Xiaohong	meiyou	chi	shenme	shucai
	Xiaohong	not	eat	what	vegetable
(18)	Xiaohong	meiyou	wei	shenme	dongwu
	Xiaohong	not	feed	what	animal

References

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language, 38*, 419–439.

Arnold, J. E., Brown-Schmidt, S., & Trueswell, J. (2007). Children's use of gender and order-of-mention during pronoun comprehension. *Language and Cognitive Processes, 22*, 527–565.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modelling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390–412.

Barr, D. J. (2008). Analyzing 'visual world' eyetracking data using multilevel logistic regression. *Journal of Memory and Language, 59*, 457–474.

Brandt-Kobele, O.-C., & Höhle, B. (2010). What asymmetries within comprehension reveal about asymmetries between comprehension and production: The case of verb inflection in language acquisition. *Lingua, 120*, 1910–1925.

Cheng, L. L.-S. (1994). Wh-words as polarity items. *Chinese Languages and Linguistics, 2*, 615–640.

Choi, Y., & Mazuka, R. (2003). Young children's use of prosody in sentence processing. *Journal of Psycholinguistic Research, 32*, 197–217.

Choi, Y., & Trueswell, J. C. (2010). Children's (in)ability to recover from garden paths in a verb-final language: Evidence for developing control in sentence processing. *Journal of Experimental Child Psychology, 106*, 41–61.

Clackson, K., Felser, C., & Clahsen, H. (2011). Children's processing of reflexives and pronouns in English: Evidence from eye-movements during listening. *Journal of Memory and Language, 65*, 128–144.

Cooper, R. M. (1974). The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology, 6*, 84–107.

Cutler, A., Dahan, D., & van Donselaar, W. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and Speech, 40*, 141–201.

Ferreira, F., Anes, M. D., & Horine, M. D. (1996). Exploring the use of prosody during language comprehension using auditory moving window technique. *Journal of Psycholinguistic Research, 25*, 273–290.

Grassmann, S., & Tomasello, M. (2007). Two-year-olds use primary sentence accent to learn new words. *Journal of Child Language, 34*, 677–687.

Grassmann, S., & Tomasello, M. (2010). Prosodic stress on a word directs 24-month-olds' attention to a contextually new referent. *Journal of Pragmatics, 42*, 3098–3105.

Höhle, B., Berger, F., Müller, A., Schmitz, M., & Weissenborn, J. (2009). Focus particles in children's language: Production and comprehension of *auch* 'also' in German learners from 1 year to 4 years of age. *Language Acquisition, 16*, 36–66.

Huang, C.-T. J. (1982a). Move *wh* in a language without *wh*-movement. *The Linguistic Review, 1*, 369–416.

Huang, C.-T. J. (1982b). *Logical relations in Chinese and the theory of grammar*. MIT: Doctoral dissertation.

Huang, C.-T. J., Li, Y.-H. A., & Li, Y. F. (2009). *The syntax of Chinese*. Cambridge: Cambridge University Press.

Ito, K., Jincho, N., Minai, U., Yamane, N., & Mazuka, R. (2012). Intonation facilitates contrast resolution: Evidence from Japanese adults and 6-year olds. *Journal of Memory and Language, 66*, 265–284.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language, 59*, 434–446.

Marslen-Wilson, W. D., Tyler, L. K., Warren, P., Grenier, P., & Lee, C. S. (1992). Prosodic effects in minimal attachment. *Quarterly Journal of Experimental Psychology, 45*, 73–87.

Nagel, H. N., Shapiro, L. P., Tuller, B., & Nawy, R. (1996). Prosodic influences on the resolution of temporary ambiguity during on-line sentence processing. *Journal of Psycholinguistic Research, 25*, 319–344.

R Development Core Team (2011). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org>.

Sedivy, J., Tanenhaus, M., Chambers, C., & Carlson, G. (1999). Achieving incremental semantic interpretation through contextual representation. *Cognition, 71*, 109–147.

Sekerina, I.A., & Trueswell, J.C. (in press). Interactive processing of contrastive expressions by Russian children. *First Language.*

Sekerina, I. A., Stromswold, K., & Hestvik, A. (2004). How do adults and children process referentially ambiguous pronouns? *Journal of Child Language, 31*, 123–152.

Snedeker, J., & Trueswell, J.C. (2001). Unheeded cues: Prosody and syntactic ambiguity in mother–child communication. In *Paper presented at the 26th Boston University conference on language development, Boston, MA*.

Snedeker, J., & Trueswell, J. C. (2003). Using prosody to avoid ambiguity: Effects of speaker awareness and referential context. *Journal of Memory and Language, 48*, 103–130.

Snedeker, J., & Trueswell, J. C. (2004). The developing constraints on parsing decisions: The role of lexical-biases and referential scenes in child and adult sentence processing. *Cognitive Psychology, 49*, 238–299.

Snedeker, J., & Yuan, S. (2008). Effects of prosodic and lexical constraints on parsing in young children (and adults). *Journal of Memory and Language, 58*, 574–608.

Speer, S. R., Kjelgaard, M. M., & Dobroth, K. M. (1996). The influence of prosodic structure on the resolution of temporary syntactic closure ambiguities. *Journal of Psycholinguistic Research, 25*, 249–271.

Tanenhaus, M., Spivey-Knowlton, M., Eberhard, K., & Sedivy, J. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science, 268*, 1632–1634.

Trueswell, J. C., Sekerina, I., Hill, N. M., & Logrip, M. L. (1999). The kindergarten-path effect: Studying on-line sentence processing in young children. *Cognition, 73*, 89–134.

Verschueren, J. (1999). *Understanding pragmatics*. London: Edward Arnold.

Warren, P., Grabe, E., & Nolan, F. (1995). Prosody, phonology and parsing in closure ambiguities. *Language and Cognitive Processes, 10*, 457–486.

Weber, A., Grice, M., & Crocker, M. W. (2006). The role of prosody in the interpretation of structural ambiguities: A study of anticipatory eye movements. *Cognition, 99*, B63–B72.

Wells, B., Peppé, S., & Goulandris, N. (2004). Intonation development from five to thirteen. *Journal of Child Language, 31*, 749–778.

Zhou, P., & Crain, S. (2011). Children's knowledge of the quantifier *dou* in Mandarin Chinese. *Journal of Psycholinguistic Research, 40*, 155–176.

Zhou, P., Su, Y., Crain, S., Gao, L. Q., & Zhan, L. K. (in press). Children's use of phonological information in ambiguity resolution: A view from Mandarin Chinese. *Journal of Child Language*.