Children's use of phonological information in ambiguity resolution: a view from Mandarin Chinese

PENG ZHOU, YI (ESTHER) SU, STEPHEN CRAIN, LIQUN GAO and LIKAN ZHAN

Journal of Child Language / Volume 39 / Issue 04 / September 2012, pp 687 - 730
DOI: 10.1017/S0305000911000249, Published online: 14 September 2011

J. Child Lang. **39** (2012), 687–730. © Cambridge University Press 2011
doi:10.1017/S0305000911000249

Children's use of phonological information in ambiguity resolution: a view from Mandarin Chinese*

PENG ZHOU, YI (ESTHER) SU AND STEPHEN CRAIN
Macquarie University

AND

LIQUN GAO AND LIKAN ZHAN
Beijing Language and Culture University

(*Received* 27 *September* 2010 – *Revised* 24 *January* 2011 – *Accepted* 3 *June* 2011 – *First published online* 14 *September* 2011)

ABSTRACT

How do children develop the mapping between prosody and other levels of linguistic knowledge? This question has received considerable attention in child language research. In the present study two experiments were conducted to investigate four- to five-year-old Mandarin-speaking children's sensitivity to prosody in ambiguity resolution. Experiment 1 used eye-tracking to assess children's use of stress in resolving structural ambiguities. Experiment 2 took advantage of special properties of Mandarin to investigate whether children can use intonational cues to resolve ambiguities involving speech acts. The results of our experiments show that children's use of prosodic information in ambiguity resolution varies depending on the type of ambiguity involved. Children can use prosodic information more effectively to resolve speech act ambiguities than to resolve structural ambiguities. This finding suggests that the mapping between prosody and semantics/pragmatics in young children is better established than the mapping between prosody and syntax.

[*] This research was partially supported by the Macquarie Centre for Cognitive Science Postgraduate Grant to Peng Zhou, and the Chinese National Social Science Grant 09BYY022 to Liqun Gao. We are grateful to two anonymous reviewers and the action editor for their insightful comments and suggestions on an earlier version of the article. We would also like to thank the children and adults who took part in the experiments for their time and patience. Address for correspondence: Peng Zhou, ARC Centre of Excellence in Cognition and its Disorders, Macquarie University, Sydney, NSW 2109, Australia. tel: +61 2 9850 6735; fax: +61 2 9850 6059; e-mail address: peng.zhou@mq.edu.au

INTRODUCTION

One of the main goals of research in child language is to understand how children develop the mapping between phonology and other levels of linguistic knowledge (e.g. syntax, semantics and pragmatics). Understanding children's mastery of the interface between phonology and other parts of the language apparatus is crucial for modelling the human language processing system, for example in understanding the role prosody plays in sentence production and comprehension. There have been a few production studies of children's use of prosodic information, and the conclusion seems to be that children are as proficient as adults in their use of prosody in production (Cutler & Swinney, 1987; Hornby & Hass, 1970; MacWhinney & Bates, 1978; Wells, Peppé & Goulandris, 2004). For example, a study of four-year-old children's productions revealed adult-like use of stress[1] in describing sequences of pictures (Hornby & Hass, 1970). Further confirmation is summarized in Cutler and Swinney (1987).

Much previous research has assessed children's use of prosodic information in sentence comprehension. These studies have yielded mixed results. In an extended review of the use of prosody in first language acquisition, Speer and Ito (2009) concluded that children perform poorly in comprehension. Young children have been found to differ from older children and adults in three areas of research: (a) in establishing phrasal groupings (Choi & Mazuka, 2003; Halbert, Crain, Shankweiler & Woodams, 1995); (b) in resolving ambiguities of phrasal attachment (Snedeker & Trueswell, 2001; Snedeker & Yuan, 2008); and (c) in associating focus particles with appropriate expressions (Gualmini, Maciukaite & Crain, 2002; Halbert *et al.*, 1995; Höhle, Berger, Müller, Schmitz & Weissenborn, 2009; Hüttner, Drenhaus, van de Vijver & Weissenborn, 2004; Szendröi, 2004). Taken together, the findings of previous research invite the conclusion that, although young children process prosodic information, and use prosodic information correctly in their productions, they are able to use such information much less effectively than older children and adults in deciding on the intended interpretation of sentences spoken to them.

Maratsos (1973) found poor performance by three- and four-year-old children in understanding the role of stress in fixing the reference of stressed pronouns, and Solan (1980) found that even older children experienced difficulties in fixing the reference of stressed pronouns. This asymmetry in production versus comprehension is paradoxical, since in most other aspects of language acquisition, comprehension far outstrips production.

Cutler and Swinney (1987) suggested that the paradox could be resolved if sensitive on-line measures were used in studies of children's knowledge of

[1] Stress refers to the relative prominence of a particular syllable, word or phrase in a certain prosodic structure. Throughout the article, stress is indicated by capital letters.

the interface relations between phonology and syntax in comprehension. This prediction was only partially confirmed, however. On-line studies (such as eye-movement recording) have revealed children's sensitivity to phonological cues in language comprehension, as compared to off-line measures (e.g. Snedeker & Yuan, 2008; Höhle *et al.*, 2009). But even studies using on-line measures have not consistently found that children are able to utilize these cues in making decisions about which interpretations to keep, and which to discard, for a majority of linguistic constructions.

As noted, negative conclusions about children's use of prosody in comprehension have been obtained in experimental investigations of children's use of stress to identify the linguistic expressions that speakers associate with focus particles. The earliest study of this was by Halbert *et al.* (1995), who investigated three- to six-year-old English-speaking children's sensitivity to prosodic cues in understanding ambiguous sentences with the focus particle *only*. Consider sentence (1), for example.

(1) Bill only gave a banana to Sue.

In this sentence, the focus particle *only* can be associated either with the direct object *a banana* or with the indirect object *Sue*. These different associations yield different truth conditions. Prosodic cues can mark the speaker's intended association. For example, with a stress on *a banana*, as in *Bill only gave A BANANA to Sue*, the sentence conveys the meaning that Bill gave a banana, but nothing else, to Sue. On the other hand, with a stress on *to Sue*, as in *Bill only gave a banana TO SUE*, the sentence conveys the meaning that Bill gave Sue, but no one else, a banana.

In the Halbert *et al.* (1995) study, a group of adults were interviewed. The adults proved to be extremely proficient in using stress to identify the intended interpretations of sentences like (1). However, only thirteen of the thirty-one three- to six-year old children performed at the same level as adults. The other eighteen children tested were apparently unable to use stress at all. These children ignored the stress in determining the interpretations they assigned to sentences. They interpreted all of the test sentences using a default strategy in which the focus particle *only* was consistently associated with either the direct object or the indirect object. Similar results for four- to five-year-old English-speaking children were obtained by Gualmini *et al.* (2002). The finding that children are not sensitive to stress in identifying the linguistic expressions associated with focus particles has been replicated for German-speaking children, using the focus particle *auch* (Hüttner *et al.*, 2004), and for Dutch-speaking children, using the focus particle *alleen* (Szendröi, 2004). A study of European Portuguese by Costa and Szendröi (2006) found that stress did not suffice for correct comprehension by three- to five-year-olds, whereas syntactic

cues (i.e. scrambling) led to adult-like performance in resolving ambiguities involving focus particles.

A more positive result was reported in a recent study by Höhle *et al.* (2009). These researchers used an on-line technique, eye-tracking, to investigate how German-speaking children interpret sentences with the accented and unaccented focus particle *auch*, and found that children exhibited adult-like comprehension, in contrast to previous studies like Hüttner *et al.* (2004).

Another sentence-level phenomenon involving prosody is the construction of constituents. Consider the examples in (2).

(2) a. Bill gave/threw her cat food.
b. Bill gave/threw her cat FOOD.
c. Bill gave/threw her CAT food.

As (2) illustrates, sequences of words with verbs like *give* and *throw* are sometimes ambiguous between a double object construction where the cat was given food, or a compound noun construction where some woman or girl was given cat food. In English, stress is used to disambiguate, so (2b) is analyzed as a double object construction, and (2c) as a compound noun construction. Halbert *et al.* (1995) conducted an experiment with eighteen children between the ages of 3;0 and 5;3 to see if they used stress as the basis for deciding how sequences of words were grouped together. Children's success in using stress was verb-specific. Most children used stress as a cue to assign the appropriate phrasal structure with the verb *throw*, but less than half were successful with the verb *give*. A subsequent study by Choi and Mazuka (2003) also found that young Korean-speaking children were unable to use prosody to resolve structural ambiguities (what they call phrasal grouping ambiguities). It is worth noting that children in the Choi and Mazuka study were able to use prosody to resolve ambiguities at the word level, i.e. word-segmentation ambiguities.

Another sentence-level ambiguity can often be resolved using prosodic information, namely the attachment of prepositional phrases. Snedeker and Trueswell (2001) investigated whether four- to five-year-old English-speaking children can use prosodic cues to disambiguate sentences like (3).

(3) Tap the frog with the flower.

There are two possible attachments of the preposition phrase *with the flower* in (3). It can be associated either with the verb *tap* or with the noun phrase *the frog*. When it is associated with the verb *tap*, the sentence conveys the instruction to 'tap the frog using the flower', the instrument reading. When it is associated with the noun phrase *the frog*, the sentence conveys the instruction to 'tap the frog that has the flower', the modifier reading. Prosodic cues can be used to distinguish between these two interpretations.

A prosodic boundary between *frog* and *with*, as in [*tap the frog*] [*with the flower*], encourages the instrument reading. By contrast, a prosodic boundary between *tap* and *the*, as in [*tap*] [*the frog with the flower*], encourages the modifier reading. In the Snedeker and Trueswell study, children were presented auditorily with either of these two prosodic versions of the test sentences and they were asked to act upon objects in the experimental workspace, based on their interpretation of the sentences. Children did not respond differently to the two versions of the test sentences.

Snedeker and Yuan (2008) hypothesized that children's failure to use prosodic cues to resolve ambiguities might have been masked, at least in part, by features of the experimental design. In the Snedeker and Trueswell (2001) study, for example, prosody was manipulated within subjects, so each child listened to both versions of the test sentences in the same experimental session. Snedeker and Yuan 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. To investigate this possibility, the Snedeker and Yuan study used a blocked experimental design in which half of the subjects heard the instrument version of the sentences (i.e. with a prosodic boundary between *frog* and *with*, like [*You can feel the frog*] [*with the feather*]) in the first session of the experiment, and then heard the modifier version (i.e. with a prosodic boundary between *feel* and *the*, as in [*You can feel*] [*the frog with the feather*]) in the second session. For the other half of the children, the sentences were presented in reverse order. As in the Snedeker and Trueswell study, the task was to act upon objects based on the spoken sentences, and children's eye-movements were recorded. The finding was that children responded differently to the two versions of the test sentences and they looked more to the intended object based on the prosodic information provided. Children were able to use the prosodic boundary as a cue in selecting the intended interpretation.

Another linguistic phenomenon in which prosody plays a role is in contrastive meanings. For example, when adults hear the adjective + noun expression 'red butterfly' with a stress on the adjective, as in *RED butterfly*, they infer the existence of at least one non-red butterfly in the context. By contrast, when the stress is 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. An eye-tracking study by Sekerina and Trueswell (in press) revealed that Russian-speaking children used the prosodic cues of stress on the adjective (as in *RED butterfly*) to facilitate identification of the intended referent, but only when the adjective and noun were adjacent. If a verb intervened between the adjective and the noun, an early contrastive stress did not facilitate the identification of the intended referent.

It seems that children's knowledge of the role of prosody in sentence comprehension is more likely to be revealed in studies using on-line

methodologies like eye-tracking. As compared to off-line methodologies, like picture-selection and judgement tasks, eye-tracking is a more sensitive testing paradigm to demonstrate children's comprehension abilities. The recording of children's eye-movements during spoken language comprehension offers a fairly direct, real-time indication of the child's attentional state during spoken language comprehension. Under most natural circumstances, one can assume that where the child is looking reflects what he or she views as relevant to the task and to the ongoing comprehension process (Trueswell, 2008; Trueswell & Gleitman, 2007). The basic premise behind this paradigm, as discussed by Tanenhaus (e.g. Tanenhaus, 2007), is that by measuring how visual–attentional states line up in time with the successive arrival of words and phrases, researchers can gain insight into the real-time processes by which listeners organize utterances structurally and semantically, and how they map these representations onto the events and objects they denote. To accept this link between eye-movement and interpretation, one need only believe that, to a useful approximation, the mind is going where the eye is going (e.g. Trueswell & Gleitman, 2007).

Therefore, the first experiment in the present study used an eye-tracking methodology to investigate the extent to which four- to five-year-old Mandarin-speaking children use stress to resolve ambiguities related to focus particles. Given the fact that previous off-line judgement tasks failed to find a prosodic effect in children's resolution of structural ambiguities, we wanted to use this on-line task to maximize the chances of detecting Mandarin-speaking children's sensitivity to stress in resolving structural ambiguities. The second experiment used a Question–Statement task to look at children's use of prosody in resolving another type of ambiguity, namely ambiguities involving different speech acts (i.e. asking a question versus making a statement). Since this ambiguity involves the pragmatic use of prosody, i.e. prosody is used to distinguish between two basic communicative acts, we felt that it would be more natural to test children in an act-out context. In addition, previous studies of children's development of pragmatic knowledge show that children from 2;6 to 3;0 easily comprehend the illocutionary force of utterances (i.e. whether the utterance expresses a request, a *yes/no* question, an offer and so on) (Bara & Bucciarelli, 1998; Bernicot & Legros, 1987; Bucciarelli, Colle & Bara, 2003; Reeder, 1980). Given this finding, we believe that children's sensitivity to prosody in resolving this type of ambiguity should be detected, even using an off-line judgement task. The two experiments reported here comprise the first investigations of Mandarin-speaking children's use of prosodic information in sentence comprehension. In the next section we discuss the kind of ambiguity investigated in Experiment 1.

HOW STRESS IS USED TO RESOLVE SYNTACTIC AMBIGUITIES IN MANDARIN

In Mandarin Chinese, the acoustic correlates of stress are mainly the duration and the high point of the pitch contour of the critical words. It is realized by expanding the pitch range and duration of the accented words (Garding, 1987; Jin, 1996; Wang, Lü & Yang, 2002; Xu, 1999; Zhong, Wang & Yang, 2001).

One characteristic of focus particles is that they associate with a unique expression in a sentence. This expression is called the focus element. Syntactically, the focus particle *zhiyou* 'only' can only associate with elements in its c-command domain[2] (e.g. Jackendoff, 1972; Reinhart, 2004; 2006). Semantically, the interpretation of a focus expression invokes a set of alternatives to the focus element. These alternatives are presupposed, in the sense that they should have already been introduced in the discourse context (Horn, 1969; Kiss, 1998; Krifka, 1991; Rooth, 1985; 1992). Consider sentence (4), for example. The particle *de* in the sentence is a possessive marker.

(4) Zhiyou Yuehan de pingguo shi hongde.
only John DE apple is red
'Only John's apple is red.'

a. Only [John's apple]$_F$ is red.
b. Only [John]$_F$'s apple is red.

In this sentence, the focus particle *zhiyou* 'only' can associate either with the entire subject noun phrase (NP) *Yuehan de pingguo* 'John's apple', as in (4a), or with the modifier of the subject NP *Yuehan* 'John', as in (4b). There are two possible focus elements because both phrases are in the c-command domain of the focus particle. (Throughout the article, the focus element is indicated by F-brackets.) If *zhiyou* 'only' is associated with the entire subject NP *Yuehan de pingguo* 'John's apple', the sentence conveys the meaning that John's apple is red and nothing else is red, as in (5). If *zhiyou* 'only' is associated with the modifier *Yuehan* 'John', the sentence conveys the meaning that John' s apple is red and no one else's apple is red, as represented in (6).[3] As noted earlier, the contrast sets are presupposed.

[2] Two definitions of the c-command relation are generally accepted. One is Reinhart's (1976), that is, node A c-commands node B if neither A nor B dominates the other and the first branching node which dominates A dominates B; and the other is Aoun and Sportiche's (1983), namely, A c-commands B, iff every maximal projection dominating A dominates B, and A does not dominate B.

[3] Another reading is possible in English, though not in Mandarin Chinese. On this reading the only thing that belongs to John and is red is his apple. Adult speakers of Mandarin Chinese do not assign this reading to the sentence.

(5) John's apple is red, and nothing else (in the discourse context) is red.

(6) John's apple is red, and no one else's apple (in the discourse context) is red.

For our purposes, the point is that prosody provides the necessary information to disambiguate between the two interpretations. Specifically, a stress on the head noun *pingguo* 'apple', as in (7), encourages the interpretation in (5); and a stress on the modifier *Yuehan* 'John', as in (8), encourages the interpretation in (6).[4]

(7) Zhiyou Yuehan de PINGGUO shi hongde.
only John DE apple is red
'Only John's APPLE is red.'

(8) Zhiyou YUEHAN de pingguo shi hongde.
only John DE apple is red
'Only JOHN's apple is red.'

The first experiment was designed to see whether Mandarin-speaking children use stress to distinguish between these two interpretations.

EXPERIMENT 1

METHOD

Participants

We tested forty-four monolingual Mandarin-speaking children between the ages of 4;1 and 4;10 (mean = 4;5). 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. In addition, thirty-eight Mandarin-speaking adults were tested for comparison. All of the adult controls were students at Beijing Language and Culture University. They had no self-reported speech or hearing disorders. They ranged in age from 23 to 27, with a mean age of 25.

[4] We conducted a survey of twenty Mandarin-speaking adults to see what their interpretation of sentences like (4) is without stress. In the interview, subjects read sentences like (4) in its written form and were then asked to tell us the meanings of these sentences. The finding was that all the twenty adults interpreted sentences like (4) as having the meaning where the focus particle is associated with the modifier. For example, we found that all twenty adults interpreted sentence (4) as having the meaning that John is the only person whose apple is red. In other words, subjects interpreted the sentences without stress as having the same meaning as their counterparts with a stress on the modifier. Based on this finding, we know that the default reading of Experiment 1 sentences is the one in which the focus particle is associated with the modifier.

Procedures and materials

Both children and adults were tested using the visual world paradigm (Tanenhaus, Spivey-Knowlton, Eberhard & Sedivy, 1995). Two versions of this paradigm have to date been successfully used with children, one in which participants were asked to act upon objects based on spoken instructions (e.g. Trueswell, Sekerina, Hill & Logrip, 1999; Snedeker & Trueswell, 2004; Snedeker & Yuan, 2008) and the other in which participants were asked to decide whether a spoken sentence accurately described a visually co-present picture (e.g. Arnold, Brown-Schmidt & Trueswell, 2007). In this study we adopted the second version. 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 sampling rate was 500 Hz. The picture stimuli were displayed on the monitor. Spoken test sentences were presented to the participants through the PC connected to two external speakers. Though the eye-tracker doesn't require head stabilization, the child participants were still held 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 picture stimuli were always about two characters, Xiaoming and 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 Chinese. The two characters also had stereotypical boy and girl appearances, for example, the girl character Xiaohong wears her hair in two braids, as in Figure 1.

In the pictures, both characters had two objects. One character had two objects that neither shared form nor colour (target character), and the other character had the same objects as the target character, but the two objects matched in colour (contrastive character). In Figure 1, for example, both Xiaoming (left) and Xiaohong (right) had scissors and a clock. Xiaoming's scissors and clock were different in colour (i.e. green scissors and yellow clock), but Xiaohong's scissors and clock were the same colour (i.e. yellow). So in the example, Xiaoming is the target character and Xiaohong is the contrastive character.

The test sentences were always about the target character. In the test sentences, the referent of the subject phrase was the object that the target character had that matched in colour with the two objects that the contrastive character had (e.g. the yellow clock of Xiaoming). All the object phrases used in the test stimuli are disyllabic in Mandarin Chinese. For example, the corresponding sentence to the example picture stimulus is (9).

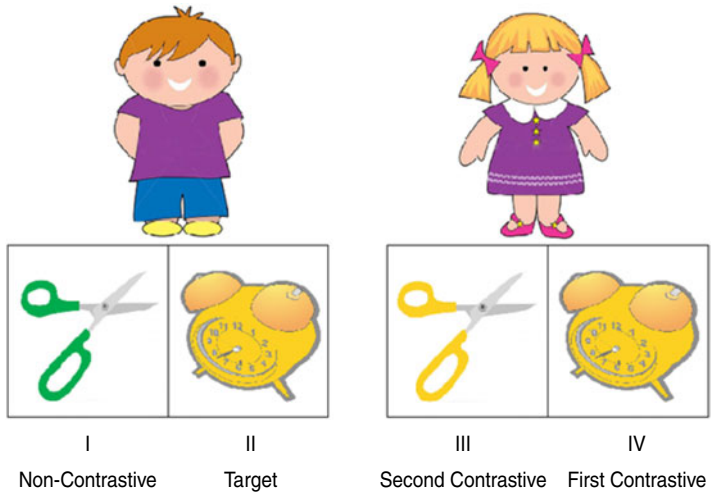
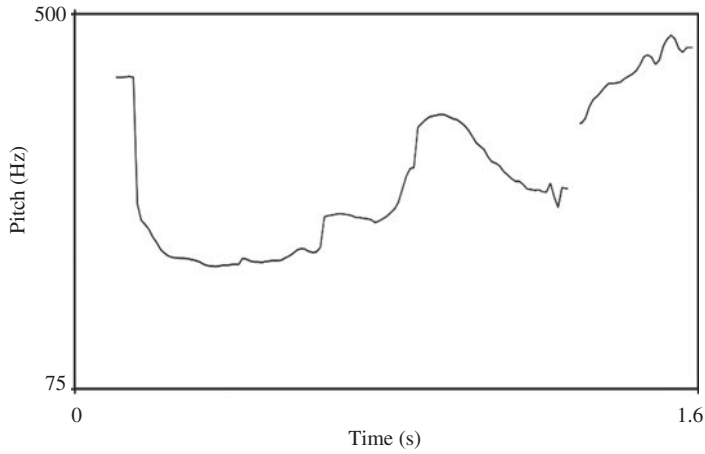


Fig. 1. (Colour online at http://journals.cambridge.org/jcl) Example of picture stimuli presented to participants (boy character on the left).

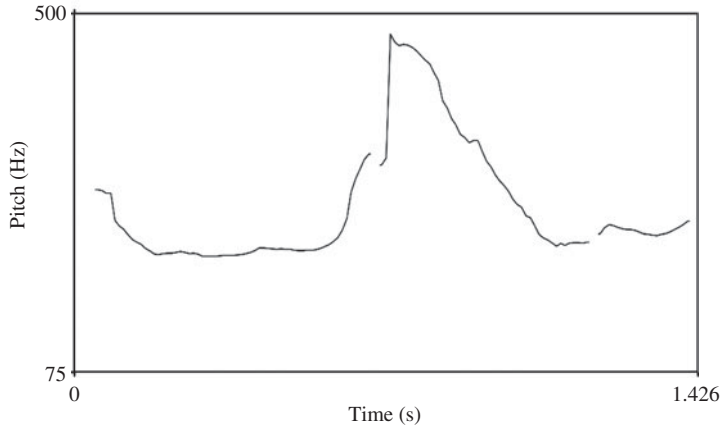
(9) Zhiyou Xiaoming de naozhong shi huangsede.
only Xiaoming DE clock is yellow
'Only Xiaoming's clock is yellow.'

(a) Xiaoming's clock is yellow; nothing else (in the discourse context) is yellow.
(b) Xiaoming's clock is yellow; no one else's clock (in the discourse context) is yellow.

As discussed above, sentences of this type are ambiguous. In sentence (9), for example, the focus particle *zhiyou* 'only' can be associated either with the subject NP (i.e. *Xiaoming de naozhong* 'Xiaoming's clock') or the modifier of the subject NP (i.e. *Xiaoming* 'Xiaoming'), thereby yielding two interpretations as in (9a) and (9b). But prosodic cues can be used to distinguish between the two readings. To be specific, a stress on the head noun (i.e. *naozhong* 'clock'), as in *Zhiyou Xiaoming de NAOZHONG shi huangsede* 'Only Xiaoming's CLOCK is yellow', encourages the interpretation in (9a), and a stress on the modifier (i.e. *Xiaoming* 'Xiaoming'), as in *Zhiyou XIAOMING de naozhong shi huangsede* 'Only XIAOMING's clock is yellow', encourages the interpretation in (9b). The pitch contours for *Xiaoming de NAOZHONG* 'Xiaoming's CLOCK' versus *XIAOMING de naozhong* 'XIAOMING's clock' are illustrated in Figure 2: stress occurs later in *Xiaoming de NAOZHONG* 'Xiaoming's CLOCK' than in *XIAOMING de naozhong* 'XIAOMING's clock'.



Xiaoming de NAOZHONG 'Xiaoming's CLOCK'



XIAOMING de naozhong 'XIAOMING's clock'

Fig. 2. Pitch contours for the noun phrase *Xiaoming de naozhong* 'Xiaoming's clock' with a stress on the head noun *Xiaoming de NAOZHONG* (upper panel) and with a stress on the modifier *XIAOMING de Naozhong* (lower panel).

Each test sentence was recorded in two prosodic versions, one with a stress on the head noun and the other with a stress on the modifier noun. We will refer to the two types of sentences as Accent-H condition and Accent-M condition respectively. In order to make the test sentences sound more natural, each test sentence was preceded by a short context. For example, sentence (9) was preceded by a context, as in (10).

TABLE 1. *Duration analyses for the spoken sentences*

Sentence segment	Mean for Accent-H condition	Mean for Accent-M condition
Focus particle (e.g. *zhiyou* 'only')	695 ms	653 ms
Modifier noun (e.g. *Xiaoming* 'Xiaoming')	505 ms	747 ms
Possessive marker DE	165 ms	151 ms
Head noun (e.g. *naozhong* 'clock')	835 ms	649 ms
BE (e.g. shi 'is')	195 ms	201 ms
Adjective (e.g. *huangsede* 'yellow')	926 ms	909 ms

(10) Xiaoming he Xiaohong chuqu mai dongxi. Tamen dou
Xiaoming and Xiaohong go buy thing they all

mai-le jiandao he naozhong. Wo kankan, zhiyou
buy-ASP scissors and clock I look only

Xiaoming de naozhong shi huangsede.
Xiaoming DE clock is yellow

'Xiaoming and Xiaohong went to buy something. They both bought scissors and a clock. Let me have a look, only Xiaoming's clock is yellow.'

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-treated booth at Beijing Language and Culture University.

In order to control for potential preferences for looking at particular displayed objects, the gender and the position of the target character were counterbalanced across trials. On half of the trials, the boy character (i.e. Xiaoming) served as the target character and on the other half, the girl character (i.e. Xiaohong) served as the target character. In addition, on half of the trials the target character appeared on the left of the picture, and on the other half the target character appeared on the right of the picture.

There were eight test trials. On each trial, participants viewed a picture like that in Figure 1 and listened to a spoken sentence like those in (9) (with a stress either on the head noun or on the modifier). Their task was to verify whether the spoken sentence was a true description of the picture. Whenever they judged the spoken sentence to be wrong, they were asked to justify their rejections. The spoken sentence started 2000 ms after the appearance of the picture stimulus. Participants' eye-movements were recorded for 6 seconds from the onset of the focus particle *zhiyou* 'only'. The mean length of the spoken sentences was 3315 ms. The mean length of each word of the two types of sentences is illustrated in Table 1 and Figure 3.

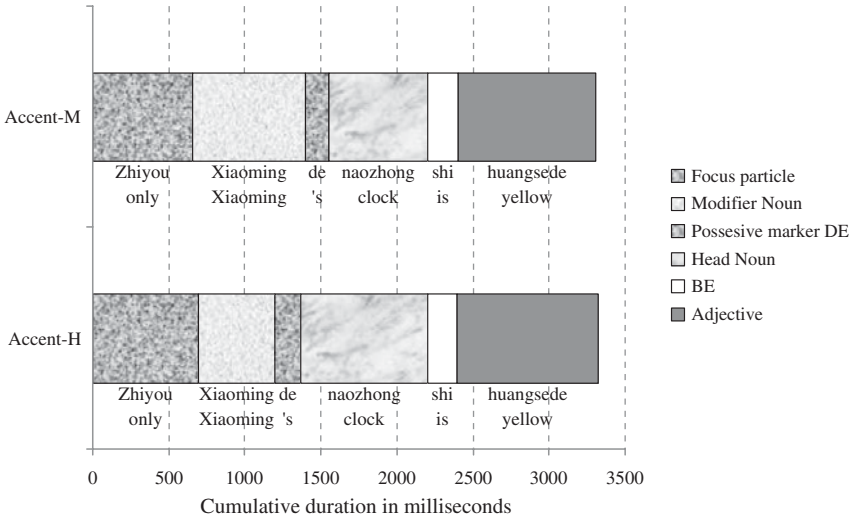


Fig. 3. Time course for the spoken sentences.

Before each trial, a picture of a little star was presented at the centre of the monitor, which anchored the beginning of each trial, and served to capture the participants' attention. This picture gave way to the trial as soon as the participant focused on the centre of the monitor.

Note that the test sentences, with either a stress on the head noun or on the modifier, were always false descriptions of the corresponding pictures. It is known that children tend to accept a test sentence to be true if they don't understand it or are unsure of the answer (Crain & Thornton, 1998). To avoid this possibility, we made all the test sentences false descriptions of the corresponding pictures. In addition, participants were asked to justify their rejections, whenever they judged the test sentences to be wrong. This was used to verify that participants reject the test sentences for the right reasons. This, however, introduced a potential problem, i.e. participants might develop an expectation that all the sentences with the focus particle *zhiyou* 'only' are false. As a consequence, participants might judge all the test sentences to be false without fully attending to the test sentences. To prevent this problem, eight filler trials were included. The picture stimuli on the filler trials were similar to those on the test trials (i.e. the gender and the position of the target character were counterbalanced across trials), and the spoken sentences on the filler trials had the same structures as the test sentences, but they were always true descriptions of the corresponding pictures. The following example is used to illustrate.

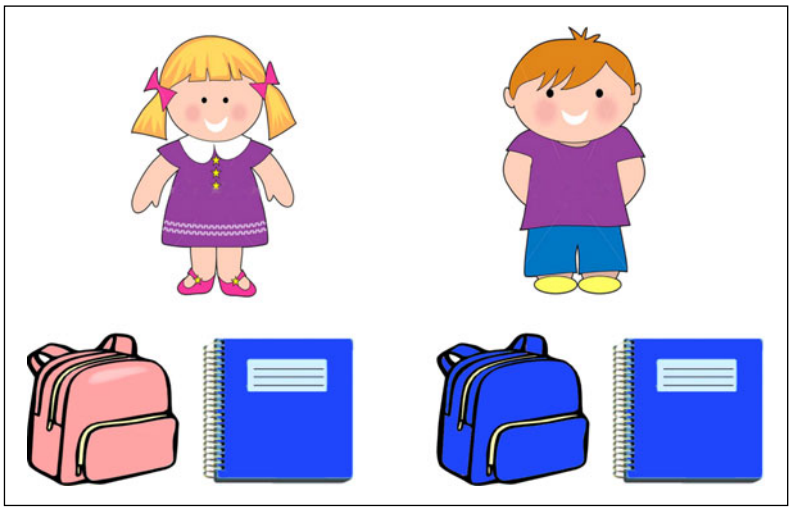


Fig. 4. (Colour online) Example of filler pictures presented to participants (girl character on the left).

(11) Zhiyou Xiaohong de shubao shi fensede.
only Xiaohong DE backpack is pink
'Only Xiaohong's backpack is pink.'

The corresponding sentence to Figure 4 is (11). Like the test sentences, each filler sentence had two prosodic versions, one with a stress on the head noun and the other with a stress on the modifier. But irrespective of where the stress was placed, the filler sentence always matched the corresponding picture. Consider sentence (11), for example. With a stress on the head noun *shubao* 'backpack', the sentence expresses the meaning that Xiaohong's backpack is pink and nothing else is pink, whereas with a stress on the modifier *Xiaohong* 'Xiaohong', the sentence expresses the meaning that Xiaohong's back pack is pink and no one else's backpack is pink. Both interpretations are true descriptions of Figure 4.

We used a between-subject design. Participants were divided into two groups. One group (22 children and 19 adults) heard the sentences with a stress on the head noun, and the other group (22 children and 19 adults) heard the same sentences with a stress on the modifier. The test and filler trials (see Appendix A) were presented to the participants in a random order. In order to familiarize children with the task, two warm-up trials were included, one in which the spoken sentence was obviously true in the picture context and the other in which the spoken sentence was obviously false in the picture context.

Before the actual experiment, we had an introduction session to familiarize children with the procedures, in which we introduced the two characters and the objects that were presented in the pictures. And since the test sentences involve colour words, we also tested children's knowledge of colours in this session. Only those children who could tell apart the different colours used in the pictures were invited to participate in the main experimental session.

RESULTS AND DISCUSSION

One child could not clearly distinguish the different colours used in our experiment, so he did not proceed to the actual test session. Five additional children and three adults were excluded because we were unable to calibrate them on the eye-tracker. The remaining thirty-eight children (20 in Accent-H condition and 18 in Accent-M condition) and thirty-five adults (18 in Accent-H condition and 17 in Accent-M condition) were included in the final analyses.

There were two sets of data: the judgement data and the eye movement data. In analyzing the eye-movement data, participants' fixations were coded in four categories: target object (the object that the target character had and that matched in colour with the two objects that the contrastive character had), first contrastive object (the object that the contrastive character had and that matched in form and colour with the target object), second contrastive object (the object that the contrastive character had and that matched only in colour with the target object) and non-contrastive object (the object that the target character had and that shared neither the form nor the colour with the target object). For example, in Figure 1, the target object is Xiaoming's yellow clock, the first contrastive object is Xiaohong's yellow clock, the second contrastive object is Xiaohong's yellow scissors and the non-contrastive object is Xiaoming's green scissors. The proportion of fixations following the onset of the focus particle *zhiyou* 'only' for each category was computed in a time window of 4200 ms. The time window was made about 1000 ms longer than the mean length of the spoken sentences, so that participants would have sufficient time to process and integrate different levels of linguistic representations when comprehending the spoken sentences.

The critical fixation area is III (the second contrastive object). Our hypothesis is that if children are sensitive to stress in resolving ambiguities involving the focus particle *zhiyou* 'only', then they should look more to the second contrastive object in the Accent-H condition than in the Accent-M condition, since with a stress on the head noun, both the first contrastive object and the second contrastive object corresponding to the target object will be evoked, whereas with a stress on the modifier, only the first

contrastive object corresponding to the target object will be evoked. Using Figure 1 and sentence (9) (repeated here as (12)) as an example,

(12) Zhiyou Xiaoming de naozhong shi huangsede.
only Xiaoming DE clock is yellow
'Only Xiaoming's clock is yellow.'

(a) Xiaoming's clock is yellow; nothing else (in the discourse context) is yellow.
(b) Xiaoming's clock is yellow; no one else's clock (in the discourse context) is yellow.

we would expect that children will look more to Xiaohong's scissors (III) in the Accent-H condition than in the Accent-M condition, because a stress on the head noun *Naozhong* 'clock' will evoke both contrastive objects – Xiaohong's scissors (III) and Xiaohong's clock (IV) – but a stress on the modifier *Xiaoming* 'Xiaoming' will only evoke the first contrastive object – Xiaohong's clock (IV). To be more specific, a stress on the head noun *Naozhong* 'clock' encourages the interpretation as in (12a), so in order to see whether this interpretation is a true description of Figure 1, children need to check whether everything else in the picture is not yellow, and it turns out that both contrastive objects are yellow: the scissors (III) and the clock (IV). A stress on the modifier *Xiaoming* 'Xiaoming' encourages the interpretation as in (12b), thus in order to verify whether this interpretation accurately describes Figure 1, children only need to check whether the contrastive character Xiaohong' clock (IV) is yellow and it turns out that it is yellow. This interpretational difference will lead to more fixations on the second contrastive object (III) in the Accent-H condition than in the Accent-M condition.

Judgement data

Thirty-eight children (20 in the Accent-H condition and 18 in the Accent-M condition) and thirty-five adults (18 in the Accent-H condition and 17 in the Accent-M condition) were included in the final analyses. All these participants responded correctly to the filler trials, i.e. they all judged the spoken sentences to be true descriptions of the corresponding pictures.

On the test trials, the test sentences, irrespective of where stress falls, were false descriptions of the corresponding pictures. So if children could use stress to resolve the ambiguity related to the focus particle *zhiyou* 'only', then they were expected to reject the test sentences in both Accent-H and Accent-M conditions, but they were expected to give different justifications for their rejections. In the Accent-H condition, where stress was placed on the head noun, children were expected to justify their rejections of the test

sentences by making reference to the fact that the two things possessed by the contrastive character also had the property mentioned in the test sentences, whereas in the Accent-M condition, where stress was placed on the modifier, they were expected to reject the test sentences on the grounds that the same thing possessed by the contrastive character also had the property mentioned in the test sentences. Here are the results.

In the Accent-H condition, adults rejected the test sentences 100% of the time (144/144 trials), and children rejected the test sentences 75% of the time (120/160 trials) and accepted them 25% of the time (40/160 trials). When examining their justifications for rejection, we found two types of rejections from adults and one type of rejection from children. Most of the adults justified their rejections by pointing out that the two things possessed by the contrastive character also had the property mentioned in the test sentences (136/144 trials; 94·44%), which we call 'focus-on-head' rejection. Only one adult pointed out that the same thing possessed by the contrastive character also had the property mentioned in the test sentences (8/144 trials; 5·56%), which we call 'focus-on-modifier' rejection. However, all the children who rejected the test sentences justified their rejections by pointing out that the same thing possessed by the contrastive character also had the property mentioned in the test sentences (120/160 trials; 75%). In other words, children only gave 'focus-on-modifier' rejections. On the example trial, 17/18 adults rejected sentence (12) by making reference to the fact that Xiaohong's scissors and clock were also yellow, and one adult rejected the sentence on the grounds that Xiaohong's clock was also yellow. However, all the children who rejected sentence (12) pointed out that Xiaohong's clock was also yellow (15/20 children). A Mann–Whitney test was used to compare the two types of rejections between children and adults, and significant differences were found between them. Children gave 'focus-on-head' rejections 0% of the time, as compared to 94·44% by adults ($Z = 5{\cdot}77$, $p < 0{\cdot}001$). Children gave 'focus-on-modifier' rejections 75% of the time, as compared to 5·56% by adults ($Z = 4{\cdot}27$, $p < 0{\cdot}001$).

In the Accent-M condition, adults rejected the test sentences 100% of the time (136/136 trials), and children rejected the test sentences 77·78% of the time (112/144 trials) and accepted them 22·22% of the time (32/144 trials). When examining their justifications for rejection, we found only one type of rejection. All the adults pointed out that the same thing possessed by the contrastive character also had the property mentioned in the test sentences (136/136 trials; 100%). Like adults, children who rejected the test sentences all pointed out that the same thing possessed by the contrastive character also had the property mentioned in the test sentences (112/144 trials; 77·78%). In other words, participants only gave 'focus-on-modifier' rejections. On the example trial, 17/17 adults justified their rejections on the grounds that Xiaohong's clock was also yellow, and children who

rejected the test sentence all pointed out that Xiaohong's clock was also yellow (14/18 children). A Mann–Whitney test showed that there was no significant difference between children and adults in their justifications of rejections. Both children and adults gave 'focus-on-modifier' rejections only (children: 77·78% vs. adults: 100%; $Z = 2{\cdot}04$, $p = 0{\cdot}10$).

When we compared the responses of the adults and children we found that adults responded differently to the test sentences in the two conditions. In the Accent-H condition, adults gave 'focus-on-head' rejections, i.e. they rejected the test sentences by making reference to the fact that the two things possessed by the contrastive character also had the property mentioned in the test sentences (e.g. 'Xiaohong's scissors and clock were also yellow'), and in the Accent-M condition, adults gave 'focus-on-modifier' rejections, i.e. they rejected the test sentences on the grounds that the same thing possessed by the contrastive character also had the property mentioned in the test sentences (e.g. 'Xiaohong's clock was also yellow'). However, children responded similarly in the two conditions. In both Accent-H and Accent-M conditions, children gave 'focus-on-modifier' rejections.

The judgement data show that when a stress was placed on the head noun, adults associated the focus particle *zhiyou* 'only' with the head noun, and when a stress was placed on the modifier, adults associated the focus particle with the modifier. However, children tended to associate the focus particle with the modifier irrespective of where the stress was placed. These findings suggest that children cannot use stress to resolve ambiguities involving the focus particle *zhiyou* 'only', in contrast to adults. However, as we discussed in the 'Introduction', off-line judgement tasks might not be sensitive enough to detect the subtle effect of stress in children's sentence comprehension. It could be the case that children are actually sensitive to stress, but they could not carry this processing sensitivity to the desired judgement response. So we looked at the eye-movement data.

Eye-movement data

The proportion of fixations following the onset of the focus particle *zhiyou* 'only' was computed in a time window of 4200 ms for the four categories: the non-contrastive object (I), the target object (II), the second contrastive object (III) and the first contrastive object (IV). This 4200 ms time window was then partitioned into twenty-one segments, each with a duration of 200 ms. The mean length of the test sentences was 3315 ms. The data of thirty-eight children (20 in the Accent-H condition and 18 in the Accent-M condition) and thirty-five adults (18 in the Accent-H condition and 17 in the Accent-M condition) were included in the final analyses.

We hypothesized that if children are sensitive to stress in resolving ambiguities involving the focus particle *zhiyou* 'only', they will look more

to the second contrastive object in the Accent-H condition than in the Accent-M condition, since with a stress on the head noun, both the first contrastive object and the second contrastive object will be evoked, whereas with a stress on the modifier, only the first contrastive object will be evoked. However, if children are not sensitive to stress in resolving ambiguities involving the focus particle *zhiyou* 'only', then they should look to the second contrastive object equally often in the Accent-H condition and in the Accent-M condition, because irrespective of where the stress was placed, they would always be associating the focus particle *zhiyou* 'only' with the modifier, as indicated by the judgement data.

Adult data

Since the critical area is the second contrastive object, we compared the proportions of fixations in each of the twenty-one segments in this area in the two conditions. No significant differences were found in the two conditions before the time span between 1600–1800 ms. Adults started to launch more fixations to the second contrastive object in the Accent-H condition than in the Accent-M condition in the time span between 1600–1800 ms, when they heard the head noun of the test sentence. This difference reached significance in the time span between 2600–2800 ms ($t(33) = 2{\cdot}39$, $p < 0{\cdot}05$) and continued in the following time spans until the sentence was completed in the time span between 3200–3400 ms ($t(33) = 3{\cdot}08$, $p < 0{\cdot}01$). Figure 5 indicates this difference.

Child data

The proportions of fixations in each of the twenty-one segments on the second contrastive object were compared between the Accent-H condition and the Accent-M condition. No significant differences were found in the two conditions before the time span between 3200–3400 ms. Children started to look more to the second contrastive object in the Accent-H condition than in the Accent-M condition in the time span between 3200–3400 ms, after the test sentence was completed. This difference reached significance in the time span between 3400–3600 ms ($t(36) = 2{\cdot}27$, $p < 0{\cdot}05$). Figure 6 indicates this difference.

A comparison between adult and child data

When children's fixation proportion pattern was compared with adults in this critical area in the two conditions, as indicated in Figures 5 and 6, we found that children and adults exhibited similar fixation patterns. Adults started to look to the second contrastive object more often in the Accent-H

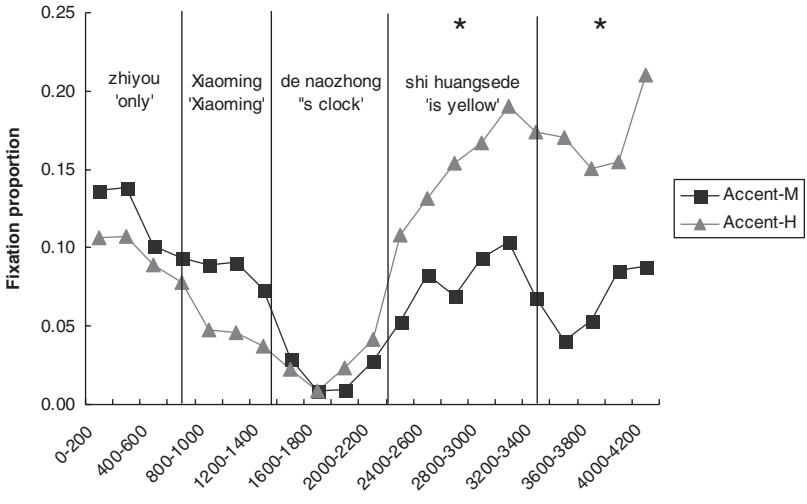


Fig. 5. Average fixation proportions over time in III (the second contrastive object) in the two conditions, adults.

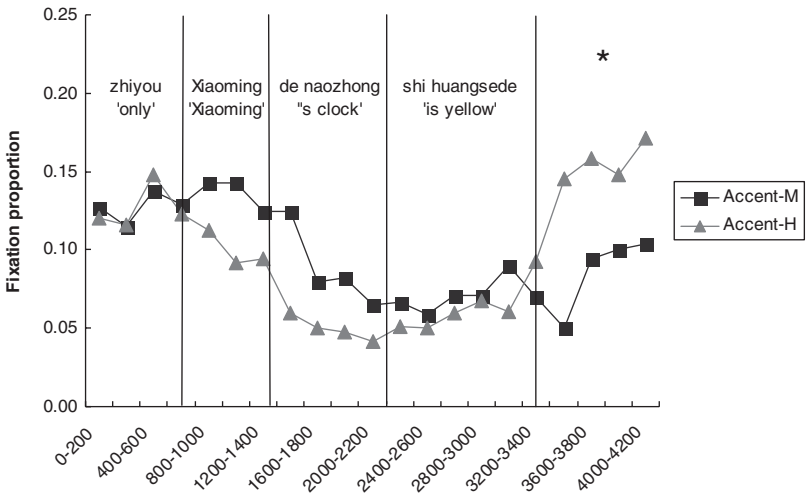


Fig. 6. Average fixation proportions over time in III (the second contrastive object) in the two conditions, children.

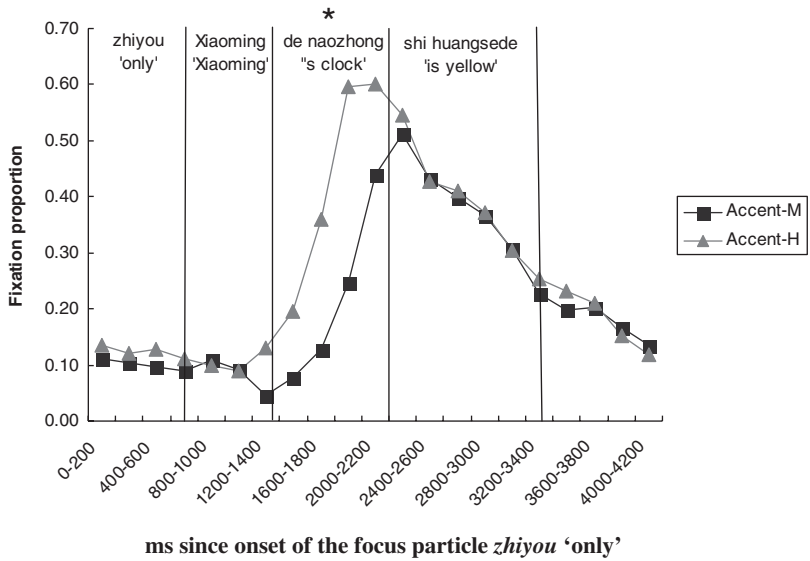


Fig. 7. Average fixation proportions over time in II (the target object) in the two conditions, adults.

condition than in the Accent-M condition in the time span between 1600–1800 ms, when they heard the head noun of the test sentence. A similar effect was found in children, but this effect was delayed. Children started to launch more fixations to the second contrastive object in the Accent-H condition as compared to the Accent-M condition in the time span between 3200–3400 ms, after the test sentence was completed.

Children and adults exhibited similar fixation patterns in the other three categories as well. Here we focus on their fixation patterns in area II (the target object). The fixation patterns in the other two categories (the non-contrastive object and the first contrastive object) are given in Appendix B. We looked in more detail at the fixation patterns in the target object area because the target object always corresponds to the head noun in the test sentences, so we expected a difference in the fixation proportions between the Accent-H condition and the Accent-M condition, if the participants are sensitive to stress. Specifically, the participants should direct their attention to the target object more often in the Accent-H condition than in the Accent-M condition after they hear the head noun, since the head noun received a stress in the Accent-H condition and thus the target object corresponding to the head noun in the Accent-H condition will be more accessible, as compared to the Accent-M condition in which the head noun did not receive a stress. This is exactly what we found. Figures 7 and 8

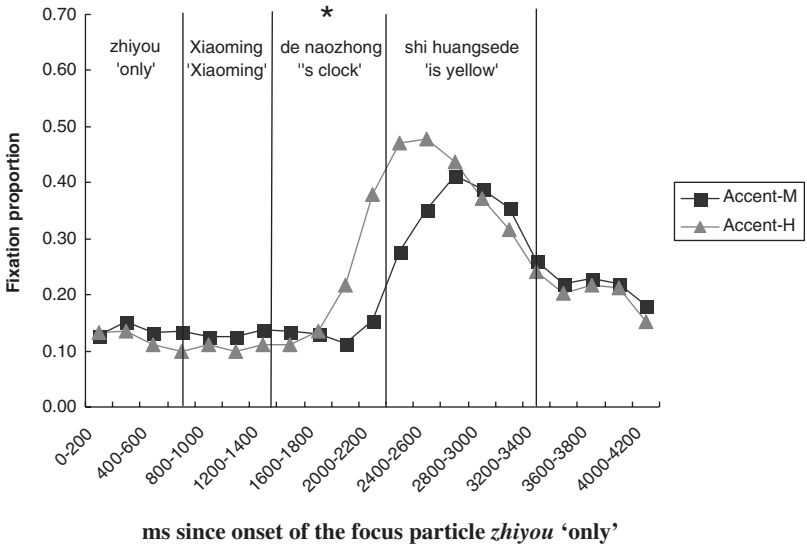


Fig. 8. Average fixation proportions over time in II (the target object) in the two conditions, children.

illustrate the average fixation proportions over time in the target object area in the two conditions by adults and children.

Adults launched more fixations to the target object in the Accent-H condition than in the Accent-M condition in the time span between 1400–1600 ms, immediately after the head noun started ($t(33)=2{\cdot}64$, $p<0{\cdot}05$). This effect was maintained in the following three time spans: 1600–1800 ms ($t(33)=4{\cdot}19$, $p<0{\cdot}001$), 1800–2000 ms ($t(33)=5{\cdot}65$, $p<0{\cdot}001$) and 2000–2200 ms ($t(33)=3{\cdot}41$, $p<0{\cdot}01$). Similar effect was found in children, but with a 400 ms delay. Children fixated more on the target object in the Accent-H condition than in the Accent-M condition in the time span between 1800–2000 ms ($t(36)=2{\cdot}91$, $p<0{\cdot}01$), after they heard the head noun. This effect was maintained in the following two time spans: 2000–2200 ms ($t(36)=4{\cdot}63$, $p<0{\cdot}001$) and 2200–2400 ms ($t(36)=3{\cdot}33$, $p<0{\cdot}01$).

The eye-movement data show that children and adults exhibited similar fixation proportion patterns when processing the test sentences. Both adults and children fixated more to the second contrastive object in the Accent-H condition than in the Accent-M condition. This finding suggests that children, like adults, are sensitive to stress in resolving ambiguities related to the focus particle *zhiyou* 'only', though the effect that was found in adults was delayed in children. Adults launched more fixations to the

second contrastive object immediately after they heard the head noun, whereas children did so after the sentence was completed. Further evidence attesting to children's sensitivity to stress comes from the eye-movement data in the target object area. Like adults, children looked more to the target object when they heard the head noun in the Accent-H condition, as compared to Accent-M condition.

DISCUSSION

The judgement data show that adults associated the focus particle *zhiyou* 'only' with the head noun when a stress was placed on the head noun, and they associated it with the modifier when a stress was placed on the modifier, whereas children tended to associate the focus particle *zhiyou* 'only' with the modifier, regardless of where the stress was. This finding seems to suggest that Mandarin-speaking children have a default reading for test sentences like (12). This is in line with the findings of our survey reported earlier indicating that the default reading of the test sentences is the one in which the focus particle is associated with the modifier.

However, when we looked at the eye-movement data, it was found that children and adults exhibited similar fixation patterns. Based on these data, we conclude that Mandarin-speaking children are actually sensitive to stress and can use it to resolve ambiguities involving the focus particle *zhiyou* 'only', though the effect that was found in adults was delayed in children.

Our findings raise two interesting questions that will be addressed in the following paragraphs. The first question is why there is a discrepancy between the judgement and the eye-tracking data, and the second is why the prosodic effect found in adults was delayed in children. Concerning the first question, one possible explanation is that this discrepancy is simply due to the different methodologies employed. In the off-line judgement task, children were asked to make a decision about whether a sentence was true or false, so the off-line judgement data offered insight into the ultimate interpretation that the child assigned to a sentence, but it provided little information about the moment-by-moment construction of the interpretation. In the eye-tracking task, however, children merely looked at the pictures while listening to sentences, and their eye-movements were recorded. This allows for a moment-by-moment record of the interpretation process, which provides a more sensitive measure of the child's sensitivity to prosody in sentence comprehension. This difference in sensitivity between the two methodologies might account for the discrepant findings. In fact, the discrepancy between off-line tasks and on-line eye-tracking tasks has been found in the acquisition of other linguistic phenomena as well.

For example, Sekerina, Stromswold and Hestvik (2004) tested children's comprehension of pronouns using both on-line and off-line tasks, and found

a dissociative pattern of performance in the two types of tasks, with eye-tracking revealing a more adult-like linguistic competence than picture-selection tasks. Höhle *et al.* (2009) conducted an eye-tracking study on children's processing of the German focus particle *auch* 'also', and found fixation patterns that indicated a higher level of competence concerning the focus particle than was found in other studies using picture-selection tasks. Brandt-Kobele and Höhle (2010) investigated children's acquisition of German verb inflections using eye-tracking and picture-selection tasks, and found discrepant patterns in the two tasks, with children successful in using their knowledge of verb inflections in sentence comprehension when tested with eye-tracking versus their failure to do so when tested via the picture-selection task. All these findings are evidence that eye-tracking is a more sensitive measure as compared to off-line methods in assessing children's comprehension abilities.

But simply attributing this discrepancy between the judgement data and the eye-tracking data to the methodological difference does not give us a complete account of children's knowledge of prosody in sentence comprehension. We still need to explain the difference between children and adults: namely, why adults exhibited the same sensitivity to stress in sentence comprehension when tested using the judgement task versus using the eye-tracking task, whereas children only demonstrated this sensitivity in the eye-tracking task, but could not carry this on-line sensitivity to the off-line judgement response.

There are at least two conceivable explanations. First, one could hypothesize that children's failure to carry this on-line sensitivity to the desired judgement response is the result of children's insufficient processing resources, like limited working memory. Along this line of reasoning, children have adult-like knowledge of prosody in sentence comprehension, i.e. they are sensitive to stress in resolving ambiguities involving the focus particle *zhiyou* 'only' as the eye-movement data suggested, but due to their limited processing resources they are not able to take the last step in incorporating this knowledge into their decision about what to say in the judgement response. Adults, by contrast, do not have problems incorporating on-line sensitivity to their judgement response, presumably because they have sufficient processing resources. Although it seems reasonable to assume that children have limited processing resources as compared to adults, this is not sufficient to conclude that these processing limitations must be the cause of children's failure to execute a desired response in the judgement task. So we propose the second possible explanation.

We attribute children's processing difficulty to particular properties of the grammar. More specifically, we propose that children's processing difficulty lies in the mapping between syntax and prosody. During the

comprehension of a sentence, listeners are engaged in the recovery of syntactic, semantic and phonological characterizations of the input. Each such characterization is computed within its own representational system (i.e. syntax, semantics and phonology) but these representations also interact with one another over time as the sentence unfolds. Adults are able to compute and integrate different levels of linguistic representations rapidly and effortlessly to arrive at an intended interpretation of the input. But we know little about how children integrate different levels of linguistic representations as they are comprehending a sentence.

Applied to our study, syntactically they are two possible interpretations for sentences like (12). The default reading associates the focus particle with the modifier. The marked reading associates the focus particle with the head noun. Phonologically, a stress is assigned to the focus element on either reading. When the focus particle is associated with the modifier, the modifier is the focus element and receives stress, and when the focus particle is associated with the head noun, the head noun receives stress. This association between stress and the focus element can be used to distinguish between the two possible interpretations. In other words, when stress is placed on the modifier, listeners should infer that the focus element is the modifier, and when stress falls on the head noun, they should infer that the focus element is the head noun. Since adults are proficient at computing and integrating different levels of linguistic knowledge, they can use the phonological information provided to disambiguate the syntactically ambiguous structure. Children, by contrast, are not as proficient as adults in mapping between phonology and syntax. Although they are sensitive to stress and are able to use it to distinguish between the two interpretations, they are not able to take the last step, which is to incorporate this on-line sensitivity into their final response. We propose that this gap between on-line sensitivity and off-line response is due to the not-yet-automatic mapping between syntax and phonology in child language.

Adults can use the mapping knowledge automatically to execute an intended response, whether stress is associated with the default reading or the marked reading. Children, however, are not as proficient as adults in using this mapping knowledge. They have no problems executing an intended response when stress is associated with the default reading (i.e. the modifier reading), but they have difficulties when stress is associated with the marked reading (i.e. the head reading). Children feel reluctant to execute a novel response by means of switching from the default reading to the marked reading, presumably because the processing of the interface knowledge between syntax and phonology is not yet automatic in child language. This might also account for the delayed prosodic effect found in children. Since the mapping between syntax and phonology in child language is not yet automatic, children are not able to use this mapping

knowledge rapidly to arrive at an intended interpretation. In other words, more processing time is needed for children before they can reach an intended interpretation.

The role of intonation in resolving ambiguities in speech acts

In a second experiment, we looked at another interface, i.e. between phonology and speech acts, to see whether children can use prosodic information in resolving ambiguities about the illocutionary force of spoken sentences. As noted earlier, Mandarin Chinese is ideally suited for evaluating children's use of intonation in resolving speech act ambiguities, since minimal sentence/intonation pairs are readily available in Chinese, but rare at best in other languages, including English. Experiment 2 takes advantage of the special properties of Mandarin, where the same sequence of words can be used to perform different speech acts. As far as we know, there have been no previous studies, in any language, looking at the role played by prosody in children's resolution of ambiguities involving speech acts.

More specifically, we investigated Mandarin-speaking children's decisions about the illocutionary force of sentences with *wh*-words. 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. Sentence (13) illustrates this.

(13) Yuehan meiyou chi shenme shuiguo.
John not eat what fruit
a. What kind of fruit didn't John eat?
b. John didn't eat any fruit.

As (13) illustrates, negative sentences with a *wh*-word like (13) are ambiguous in Mandarin Chinese.[5] This sentence can be used to pose a question, as in (13a): 'What kind of fruit didn't John eat?' Alternatively, the same sequence of words can make a statement, as in (13b): 'John didn't eat any fruit.' Intonational 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.

[5] The ambiguity of this type of sentence is due to the properties of *wh*-words in Mandarin Chinese. *Wh*-words in Mandarin Chinese can function either as question words or as indefinite NPs (roughly like English *some*). Interested readers are referred to the articles by Cheng (1991; 1994), Huang (1982), Li (1992) and Lin (1996; 1998). English has semantic ambiguities, but these are not as readily available as in Mandarin Chinese. For example, English has sentences like *Have the soldiers killed*, which can be used either to make a command *Have someone kill the soldiers*, or to pose a question *Have the soldiers killed anyone*? But it is not clear whether prosody can be used to disambiguate between these two readings, and such ambiguities are not abundant in English.

Both kinds of speech acts are used frequently in daily conversation. Experiment 2 was designed to investigate whether young Mandarin-speaking children are sensitive to intonational cues in resolving ambiguities involving different speech acts.

EXPERIMENT 2

METHOD

Participants

Thirty-eight monolingual Mandarin-speaking children participated in this experiment (mean age 4;4, range 3;6 to 4;11). They were recruited from the kindergarten at Beijing Language and Culture University and had no reported history of speech, hearing or language disorders. In addition, twenty Mandarin-speaking adults were tested as controls, all postgraduates at Beijing Language and Culture University (mean age 26, range 25 to 28). They had no self-reported speech or hearing disorders.

Procedures and materials

In order to evaluate the experimental hypothesis, we used a Question–Statement task. The experimenter acted out stories in front of the child participant using toy characters and props, and a puppet who appeared on a laptop computer screen watched the stories alongside the participant. After each story, the puppet attempted to explain to the participant what he thought had happened in the story, using a test sentence. The test sentences were prerecorded and were presented to the participant through the laptop computer connected to an external speaker to make it appear that the puppet was talking. It was made clear to the participant that the puppet did not always pay close attention to the story and thus was sometimes unsure about the outcome of the story. If that was the case, the puppet would make a guess about what happened in the story or ask the subject a question. On each trial, the participant's task was to decide whether the puppet made a statement about what happened in the story or asked a question about the story. Whenever the puppet said what happened in the story, the participant was instructed to judge whether the puppet was right or wrong. But if the puppet asked a question about the story, the participant was instructed to answer the question.

The participants were introduced to the task individually and were tested individually. In order to familiarize them with the task, they were given two practice trials before the actual test session. On one practice trial, the puppet made a statement informing the participant about what happened in the story, and on the other practice trial, the puppet asked the participant a question. Only those participants who correctly judged the puppet had

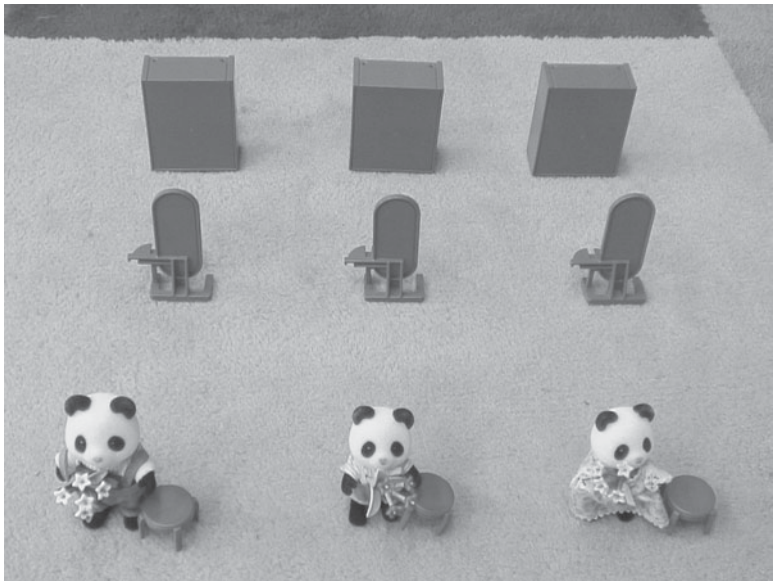


Fig. 9. The last scene of the story.

made a statement on the first trial and had posed a question on the second trial were permitted to advance to the test session. Adult controls were tested on the same task.

There were four test stories. For each story, two versions of the test sentence were created, one with a rising intonation on the *wh*-phrase and the other with a level intonation on the *wh*-phrase. The following example is used to illustrate. In a typical story, three pandas went to buy something for their new house. They came to Minnie Mouse's shop. There were four things there: three types of furniture (i.e. wardrobes, mirrors and chairs) and flowers. They first came to the wardrobes, but the wardrobes were too big for their house, so they did not buy them. Then they came to the flowers. The flowers were beautiful, so they all bought one. After that, they came to the mirrors. They wanted to buy mirrors. But when they looked into the mirrors, they found that they were not as good-looking as they expected. They thought the mirrors made them ugly. So they decided not to buy the mirrors. Finally they came to the chairs. The chairs were nice. So they all bought a chair. The last scene of the story is illustrated in Figure 9: the pandas bought the flowers and the chairs.

When the story concluded, sentence (14) was presented to the participants, either with a rising intonation on the *wh*-phrase *shenme jiaju* 'what furniture' or with a level intonation on the *wh*-phrase *shenme jiaju* 'what furniture'.

(14) Xiongmao meiyou mai shenme jiaju.
panda not buy what furniture

a. What type of furniture didn't the pandas buy?
b. The pandas didn't buy any furniture.

The intonation contours for the *wh*-phrase *shenme jiaju* 'what furniture' with a rising intonation and with a level intonation are illustrated in Figure 10.

We used a between-subject design. Participants were divided into two groups. One group (19 children and 10 adults) heard the sentences with a rising intonation on the *wh*-phrase, and the other group (19 children and 10 adults) heard the same sentences with a level intonation on the *wh*-phrase. The test sentences were produced by a female native speaker of Beijing Mandarin and she was asked to produce the test sentences in a child-directed manner. The recording was conducted in a sound-attenuated laboratory at Macquarie University.

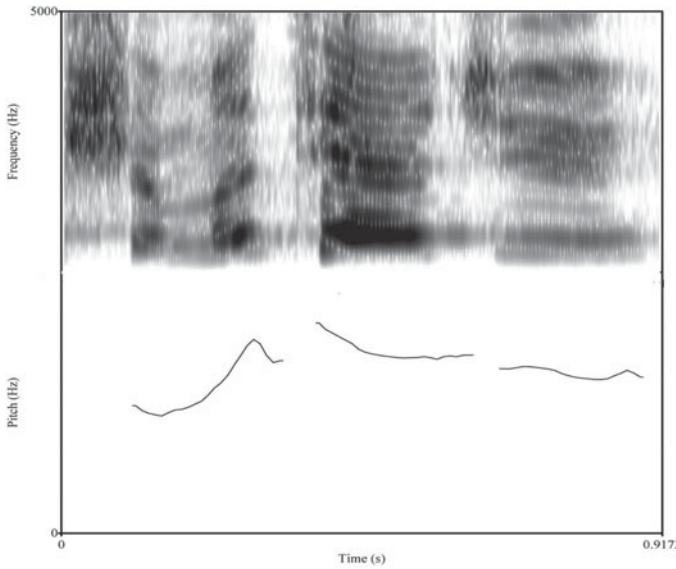
In addition to the test trials, four filler trials were included for each group. The group of participants who heard the test sentences with a rising intonation on the *wh*-phrase were presented with four negative statement fillers like (15), and the statements were all false descriptions of the corresponding stories. And the group of participants who heard the test sentences with a level intonation on the *wh*-phrase were presented with four *wh*-question fillers like (16). These fillers were used to counterbalance the question and statement interpretations throughout the trials. The test and filler trials were presented in a pseudo-random order. All the test and filler sentences are given in Appendix C.

(15) Haidao meiyou nazou jinbi.
pirate not take gold-coin
'The pirates didn't take the gold coins.'

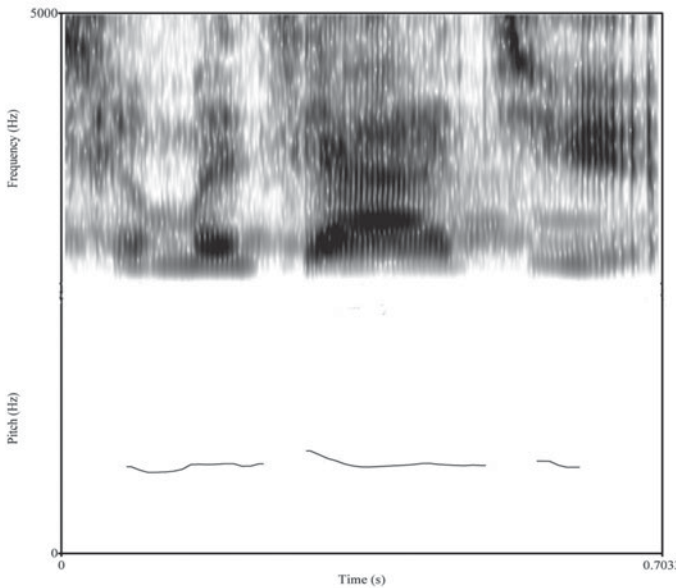
(16) Xiaoniao chi-le shenme shiwu?
bird eat-ASP what food
'What kind of food did the bird eat?'

Predictions

If children are sensitive to the intonational cues (i.e. rising intonation vs. level intonation) and can use the cues to resolve this speech act ambiguity (i.e. asking a question vs. making a statement), then they are expected to interpret sentences with a rising intonation on the *wh*-phrase as questions, and sentences with a level intonation on the *wh*-phrase as statements. On the example trial, they should respond to sentence (14) with a rising intonation by offering an answer 'wardrobes and mirrors', and they should



Shenme jiaju 'what furniture' with a rising intonation



Shenme jiaju 'what furniture' with a level intonation

Fig. 10. Intonation contours for the *wh*-phrase *shenme jiaju* 'what furniture' with a rising intonation (upper panel) and with a level intonation (lower panel).

respond to sentence (14) with a level intonation on the *wh*-phrase by rejecting the sentence on the grounds that the pandas bought the chairs.

RESULTS AND DISCUSSION

We recorded the responses of the participants to the two versions of the test sentences. All the participants responded correctly to the filler trials. So their data were all included in the final analyses.

Here are the main findings. In response to the test sentences with a rising intonation on the *wh*-phrase, children responded by providing an answer 97·37% of the time (74/76 trials) and adults did so 100% of the time (40/40 trials). A Mann–Whitney test showed that there was no significant difference in the response patterns between children and adults ($Z = 1{\cdot}05$, $p = 0{\cdot}53$). In the example story, children answered sentence (14) in one of two ways. One response, by three of the nineteen child participants, was to point to the wardrobes and the mirrors. The remaining sixteen children overtly mentioned that the pandas did not buy the wardrobes and the mirrors, saying ''Meiyou mai yichu he jingzi'' 'They didn't buy the wardrobes and the mirrors'. All the adults provided an answer ''Yichu he jingzi'' 'The wardrobes and the mirrors'.

In response to the test sentences with a level intonation on the *wh*-phrase, children responded to the test sentences by rejecting them 86·84% of the time (66/76 trials) and adults rejected them 90% of the time (36/40 trials). A Mann–Whitney test showed that there was no significant difference in the response patterns between children and adults ($Z = .38$, $p = 1$). In the example story, children and adults rejected sentence (14) by making reference to the fact that the pandas bought the chairs. Children either explicitly mentioned the chairs ('No, they bought the chairs') or they pointed to the chairs. The responses of each child on this trial are provided in Appendix D.

The results of this experiment clearly show that Mandarin-speaking children know that a rising intonation on the *wh*-phrase turns the sentences into questions, whereas a level intonation on the *wh*-phrase turns them into statements. This is compelling evidence that Mandarin-speaking children can reliably use intonational cues in resolving ambiguities involving different speech acts. When the nature of the speech act being performed depends on children's sensitivity to prosodic cues, Mandarin-speaking children are even sensitive to the fact that a level intonation indicates that the speaker is making a statement, rather than asking a question. These findings suggest that children are sensitive to intonational cues and can use them to resolve ambiguities of illocutionary force. This finding is not surprising given that this type of ambiguity involves the pragmatic use of prosody (i.e. asking a question versus making a statement). We will return to this in the 'General discussion'.

GENERAL DISCUSSION

The present study investigated the role of prosody in children's ambiguity resolution. Two experiments were conducted. Experiment 1 looked at the role of stress in children's resolution of syntactic ambiguities involving the focus particle *zhiyou* 'only', and Experiment 2 investigated the role of intonational cues in children's resolution of ambiguities involving different speech acts (i.e. asking a question vs. making a statement).

Using off-line tasks, previous research found that four- to five-year-old English-speaking children are not sensitive to prosodic cues in resolving ambiguities involving the focus particle *only* (Gualmini *et al.*, 2002; Halbert *et al.*, 1995). However, our study (Experiment 1), using an eye-tracking task, has shown that four- to five-year-old Mandarin-speaking children are sensitive to stress in resolving such ambiguities, though they are not able to take the last step in incorporating this sensitivity into their final response. We attribute children's processing difficulty to the not-yet-automatic mapping between syntax and phonology. More specifically, since the mapping between syntax and phonology in child language is not yet automatic, children are not able to use this mapping knowledge rapidly to arrive at an intended interpretation, as evidenced by the delayed prosodic effect in children's eye-movement data. The delayed effect of prosody has also been observed in some other studies investigating children's sensitivity to prosodic information in sentence comprehension (e.g. Sekerina & Trueswell, in press; Snedeker & Yuan, 2008). One reviewer raised the question of whether this delayed effect found in child language reflects a difference between children and adults in the processing of the mapping between syntax and prosody, or a difference in the whole processing system including, for example, motor control for the eye-movements. We will have to leave this question open before systematic studies are conducted to investigate the processing differences between children and adults. But as the reviewer correctly points out, it is important that we take into account both linguistic and non-linguistic factors when explaining children's non-adult behaviour.

In the second experiment, we looked at whether children are sensitive to prosodic cues in resolving ambiguities involving different speech acts. Mandarin Chinese is ideally suited for evaluating children's use of intonation in resolving speech act ambiguities, since minimal sentence/intonation pairs are readily available in Chinese, but rare at best in other languages, including English. Experiment 2 takes advantage of the special properties of Mandarin, where the same sequence of words can be used to perform different speech acts. As far as we know, there have been no previous studies, in any language, looking at the role played by prosody in children's resolution of ambiguities involving speech acts. So Experiment 2 investigated whether Mandarin-speaking children can use intonational cues

to resolve such kinds of ambiguity. The results show that four-year-old Mandarin-speaking children know that a rising intonation on the *wh*-phrase indicates a question reading, whereas a level intonation on the *wh*-phrase indicates a statement reading. Children are sensitive to intonational cues and can use them to resolve ambiguities of illocutionary force. This finding is not surprising given that this type of ambiguity involves the pragmatic use of prosody (i.e. asking a question versus making a statement). The pragmatic use of prosody should be acquired early in language development, since it is important for the development of general communication skills, e.g. making a request, asking a question, etc. In other words, the mapping between prosody and speech acts should be established early in child language.

Previous studies on children's development of speech acts (the emergence of illocutionary force) found that children from 2;6 to 3;0 can easily comprehend the illocutionary force of utterances (i.e. whether the utterance expresses a request, a *yes/no* question, an offer, and so on) (Bara & Bucciarelli, 1998; Bernicot & Legros, 1987; Bucciarelli *et al.*, 2003; Reeder, 1980). Given the previous findings, it is not surprising that children in our study can easily use intonational cues to distinguish whether a *wh*-phrase indicates a question or a statement. It would be interesting to establish the correlation between the development of prosody and the acquisition of communication skills in further studies.

Combining the data of the two experiments, we found that children can use prosodic information more effectively to resolve ambiguities of illocutionary force than to resolve syntactic ambiguities. This finding can be attributable to (i) the different methodologies used in the two experiments, (ii) the different prosodic characteristics used, or (iii) the different types of ambiguity involved.

According to the first possibility, the use of Question–Statement task might account for children's better performance in Experiment 2. More specifically, the use of toys and contexts might greatly facilitate children's comprehension of the test sentences in Experiment 2, thus children exhibited more adult-like knowledge. We want to note, first, that previous studies on English-speaking children's sensitivity to stress in resolving ambiguities related to the focus particle *only* also used judgement tasks, in which toys and contexts were used (Gualmini *et al.*, 2002; Halbert *et al.*, 1995). Children were nevertheless found not to be able to use stress to resolve this ambiguity. Thus, it seems unlikely that this difference in methodology could have resulted in the different findings in the two experiments. Two further points about the methodologies are worth making. One is that online methodologies like eye-tracking have proved to be more sensitive in detecting children's knowledge of prosodic information in sentence comprehension, as compared to off-line judgement tasks. So, if a

prosodic effect can be found in children using off-line tasks, it should also be found using the more sensitive on-line tasks. The second point relates to the reason why we used contexts and toys in Experiment 2. As mentioned in the 'Introduction', we used this task simply because the ambiguity we tested involves the pragmatic use of prosody, i.e. prosody is used to distinguish between two basic communicative acts. We felt that it would be more natural to test children's knowledge of the pragmatic use of prosody in an act-out context. Based on these considerations, it seems unlikely that the different methodologies used would have led to the different findings in the two experiments.

The second possibility would attribute the different findings to the different prosodic characteristics used in the two experiments. Since Experiment 1 used stress and Experiment 2 used intonation, it might be hypothesized that intonational cues are easier to process for children than stress. Although we don't have direct evidence as to whether children can process intonation more easily than stress, previous studies seem to suggest that young children have no difficulty in perceiving stress. For example, Grassmann and Tomasello (2007; 2010) found that children as young as two years of age can detect stress in a sentence and understand that stress is used to indicate referential newness in the sentence.

Given the results of both previous studies and our study, we attribute the discrepant findings in the two experiments to the different types of ambiguities involved. Specifically, we propose that children can use prosodic information more effectively to resolve ambiguities of illocutionary force than to resolve syntactic ambiguities, because the mapping between prosody and semantics/pragmatics is better established than the mapping between prosody and syntax in young children. As discussed earlier, children can use prosodic information more effectively to resolve ambiguities of illocutionary force, presumably because this ambiguity involves the pragmatic use of prosody. The basic idea is that the pragmatic use of prosody is crucial for the development of general communication skills, and therefore should be acquired early in language development. However, the correlation between the development of prosody and the acquisition of communication skills still remains to be empirically investigated.

An issue raised in the 'Introduction' was asymmetry between production and comprehension. It is widely acknowledged that children's production precedes comprehension in the acquisition of prosody in that children can use prosodic information correctly in their production, but they are less able to use such information than adults in deciding on the intended interpretation of sentences spoken to them (e.g. Cruttenden, 1985; Cutler & Swinney, 1987; Solan, 1980; Wells *et al.*, 2004). We wish to note that this asymmetry between comprehension and production is not limited to the acquisition of prosody. It has been found in the acquisition of other

linguistic structures as well, like English pronouns (e.g. Hendriks & Spenader, 2005/2006; Sekerina *et al.*, 2004), the German focus particle *auch* 'also' (e.g. Hüttner *et al.*, 2004; Höhle *et al.*, 2009) and verb inflection (e.g. Brandt-Kobele & Höhle, 2010; Johnson, de Villiers & Seymour, 2005). Various explanations have been put forward to account for the asymmetry (see Hendriks & Koster (2010) for a comprehensive review). We are not going into the details of these possible explanations. Instead, we want to point out that the assumption that production outstrips comprehension in these cases is mainly based on the observation that children exhibited adult-like performance in production tasks, but not in sentence comprehension tasks. So the question of how we can explain the asymmetry can be reformulated as: 'How can we account for children's non-adult behaviour in these comprehension tasks?' In the case of prosody, children's non-adult behaviour can be attributable to (i) the properties of the experimental tasks used, or (ii) the differences between children and adults in their linguistic and non-linguistic abilities.

The first possibility is based on the discrepant findings of recent eye-tracking studies and those of the earlier studies. In contrast to the earlier studies, which used off-line tasks and failed to find a prosodic effect, our eye-tracking study, together with the studies of Höhle *et al.* (2009) and Snedeker and Yuan (2008), have revealed a more adult-like linguistic competence concerning children's use of prosody in sentence comprehension. According to this possibility, if sensitive on-line measures were used, children would exhibit adult-like comprehension. In other words, the asymmetry observed in previous studies is just a methodological artefact. However, this prediction is only partially confirmed. Although on-line studies like eye-tracking can reveal a higher level of linguistic competence in children than indicated by off-line tasks, this competence is still not completely adult-like. For example, our study showed that the effect of prosody found in adults was delayed in children. This difference between children and adults still needs to be explained. This brings us to the second possibility that might account for children's non-adult performance: the differences between children and adults in their linguistic and non-linguistic abilities. For instance, in our study children's processing difficulty as indicated by the delayed effect can be attributed either to the difference between children and adults in the processing of the mapping between syntax and prosody (linguistic ability), or to a difference in the whole processing system including, for example, the motor control for the eye-movements (non-linguistic ability). The current study will have to leave this question open.

To conclude, we found that children's use of prosodic information in ambiguity resolution can vary according to the type of ambiguity involved. In our study, children can use prosodic information more effectively to

resolve ambiguities of illocutionary force than to resolve syntactic ambiguities. This finding suggests that the mapping between prosody and semantics/pragmatics might be better established than the mapping between prosody and syntax in young children. Of course, we cannot simply generalize our results to other types of syntactic ambiguities and ambiguities of illocutionary force. Further investigations of different types of syntactic ambiguities and ambiguities of illocutionary force are required. We also want to emphasize that different uses of prosody in child language should be treated differently. Some uses of prosody might be acquired earlier and thus better established than others in child language. For example, in our study the knowledge of the pragmatic use of prosody is better established than the knowledge of the use of prosody to distinguish syntactic ambiguities. In addition, we want to point out that cross-linguistic studies are necessary, if we want to understand which aspects of prosodic information and what type of ambiguities are more easily processed in child language.

REFERENCES

Aoun, J. & Sportiche, D. (1983). On the formal theory of Government. *Linguistic Review* **2**, 211–36.

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**, 527–65.

Bara, B. G. & Bucciarelli, M. (1998). Language in context: The emergence of pragmatic competence. In A. C. Quelhas & F. Pereira (eds), *Special issue of Analise Psicologica: Cognition in context*, 317–45. Lisbon: Istituto Superior de Psicologia Aplicada.

Bernicot, J. & Legros, S. (1987). Direct and indirect directives: What do young children understand? *Journal of Experimental Child Psychology* **43**, 346–58.

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**, 1910–25.

Bucciarelli, M., Colle, L. & Bara, B. G. (2003). How children comprehend speech acts and communicative gestures. *Journal of Pragmatics* **35**, 207–241.

Cheng, L. L.-S. (1991). On the typology of *wh*-questions. Unpublished doctoral dissertation, MIT.

Cheng, L. L.-S. (1994). Wh-words as polarity items. *Chinese Languages and Linguistics* **2**, 615–40.

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

Costa, J. & Szendröi, K. (2006). Acquisition of focus marking in European Portuguese: Evidence for a unified approach to focus. In V. Torrens & L. Escobar (eds), *The acquisition of syntax in Romance Languages*, 319–29. Amsterdam: John Benjamins.

Crain, S. & Thornton, R. (1998). *Investigations in Universal Grammar: A guide to experiments on the acquisition of syntax and semantics*. Cambridge, MA: MIT Press.

Cruttenden, A. (1985). Intonation comprehension in ten-year-olds. *Journal of Child Language* **12**, 643–61.

Cutler, A. & Swinney, D. A. (1987). Prosody and the development of comprehension. *Journal of Child Language* **14**, 145–67.

Garding, E. (1987). Speech act and tonal pattern in Standard Chinese. *Phonetica* **44**, 13–29.

Grassmann, S. & Tomasello, M. (2007). Two-year-olds use primary sentence accent to learn new words. *Journal of Child Language* **34**, 677–87.
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.
Gualmini, A., Maciukaite, S. & Crain, S. (2002). Children's insensitivity to contrastive stress in sentences with *only*. In S. Arunachalam, E. Kaiser & A. Williams (eds), *Proceedings of the 25th Penn Linguistics Colloquium*, 87–110. Philadelphia, PA: University of Pennsylvania.
Halbert, A., Crain, S., Shankweiler, D. & Woodams, E. (1995). Children's interpretive use of emphatic stress. Poster presented at the 8th Annual CUNY Conference on Human Sentence Processing, Tucson, AZ.
Hendriks, P. & Koster, C. (2010). Production/comprehension asymmetries in language acquisition. *Lingua* **120**, 1887–97.
Hendriks, P. & Spenader, J. (2005/2006). When production precedes comprehension: An optimization approach to the acquisition of pronouns. *Language Acquisition* **13**, 319–48.
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.
Horn, L. R. (1969). A presuppositional approach to *only* and *even*. *Proceedings of the Chicago Linguistic Society* **5**, 98–107.
Hornby, P. A. & Hass, W. A. (1970). Use of contrastive stress by preschool children. *Journal of Speech and Hearing Research* **13**, 395–99.
Huang, C.-T. J. (1982). Logical relations in Chinese and the theory of grammar. Unpublished doctoral dissertation, MIT.
Hüttner, T., Drenhaus, H., van de Vijver, R. & Weissenborn, J. (2004). The acquisition of the German focus particle *auch* 'too': Comprehension does not always precede production. *Proceedings of the 28th Annual Boston University Conference on Language Development*, on-line supplement www.bu.edu/linguistics/APPLIED/BUCLD/supp.html.
Jackendoff, R. (1972). *Semantic interpretation in generative grammar*. Cambridge, MA: MIT Press.
Jin, S. (1996). An acoustic study of sentence stress in Mandarin Chinese. Unpublished doctoral dissertation, Ohio State University.
Johnson, V. E., de Villiers, J. G. & Seymour, H. N. (2005). Agreement without understanding? The case of third person singular /s/. *First Language* **25**, 317–30.
Kiss, K. É. (1998). Identificational focus versus informational focus. *Language* **74**, 245–73.
Krifka, M. (1991). A compositional semantics for multiple focus constructions. *Cornell Working Papers in Linguistics* **10**, 127–58.
Li, Y.-H. A. (1992). Indefinite *wh* in Mandarin Chinese. *Journal of East Asian Linguistics* **1**, 125–55.
Lin, J.-W. (1996). Polarity licensing and *wh*-phrase quantification in Chinese. Unpublished doctoral dissertation, University of Massachusetts at Amherst.
Lin, J.-W. (1998). On existential polarity *wh*-phrases in Chinese. *Journal of East Asian Linguistics* **7**, 219–55.
MacWhinney, B. & Bates, E. (1978). Sentential devices for conveying givenness and newness: A cross-cultural developmental study. *Journal of Verbal Learning and Verbal Behaviour* **17**, 539–58.
Maratsos, M. P. (1973). The effects of stress on the understanding of pronominal co-reference in children. *Journal of Psycholinguistic Research* **2**, 1–8.
Reeder, K. (1980). The emergence of illocutionary skills. *Journal of Child Language* **7**, 13–28.
Reinhart, T. (1976). The syntactic domain of anaphora. Unpublished doctoral dissertation, MIT.

Reinhart, T. (2004). The processing cost of reference set computation: Acquisition of stress shift and focus. *Language Acquisition* **12**, 109–155.
Reinhart, T. (2006). *Interface strategies: Optimal and costly computations*. Cambridge, MA: MIT Press.
Rooth, M. (1985). Association with focus. Unpublished doctoral dissertation, University of Massachusetts at Amherst.
Rooth, M. (1992). A theory of focus interpretation. *Natural Language Semantics* **1**, 75–116.
Sekerina, I. A., Stromswold, K. & Hestvik, A. (2004). How do adults and children process referentially ambiguous pronouns? *Journal of Child Language* **31**, 123–52.
Sekerina, I. E. & Trueswell, J. C. (in press). Interactive processing of contrastive expressions by Russian children. *First Language*.
Snedeker, J. & Trueswell, J. C. (2001). Unheeded cues: Prosody and syntactic ambiguity in mother–child communication. Paper presented at the 26th Boston University Conference on Language Development, Boston, MA.
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**, 238–99.
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.
Solan, L. (1980). Contrastive stress and children's interpretation of pronouns. *Journal of Speech and Hearing Research* **23**, 688–98.
Speer, S. R. & Ito, K. (2009). Prosody in first language acquisition – Acquiring intonation as a tool to organize information in conversation. *Language and Linguistics Compass* **3**, 90–110.
Szendröi, K. (2004). Acquisition evidence for an interface theory of focus. In J. Van Kampen & S. Baauw (eds), *Proceedings of GALA 2003*, 457–68. Utrecht: LOT.
Tanenhaus, M. K. (2007). Spoken language comprehension: Insights from eye movements. In G. Gaskell (ed.), *Oxford handbook of psycholinguistics*, 309–326. Oxford: Oxford University Press.
Tanenhaus, M. K., Spivey-Knowlton, M., Eberhard, K. & Sedivy, J. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science* **268**, 1632–34.
Trueswell, J. C. (2008). Using eye movements as a developmental measure within psycholinguistics. In I. A. Sekerina, E. M. Fernández & H. Clahsen (eds), *Developmental psycholinguistics: On-line methods in children's language processing*, 73–96. Amsterdam: John Benjamins.
Trueswell, J. C. & Gleitman, L. R. (2007). Learning to parse and its implications for language acquisition. In G. Gaskell (ed.), *Oxford handbook of psycholinguistics*, 635–56. Oxford: Oxford University Press.
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.
Wang, B., Lü, S. N. & Yang, Y. F. (2002). The pitch movement of stressed syllable in Chinese sentences. *Acta Acustica* **27**, 234–40 [in Chinese].
Wells, B., Peppé, S. & Goulandris, N. (2004). Intonation development from five to thirteen. *Journal of Child Language* **31**, 749–78.
Xu, Y. (1999). Effects of tone and focus on the formation and alignment of Fo contours. *Journal of Phonetics* **27**, 55–105.
Zhong, X. B., Wang, B. & Yang, Y. F. (2001). The perception of prosodic word stress in standard Chinese. *Acta Psychologica Sinica* **34**, 2–9 [in Chinese].

APPENDIX A: TEST SENTENCES IN EXPERIMENT 1

Eight test sentences:

(1) Zhiyou Xiaoming de naozhong shi huangsede.
only Xiaoming DE clock is yellow
'Only Xiaoming's clock is yellow.'

(2) Zhiyou Xiaohong de yazi shi huangsede.
only Xiaohong DE duck is yellow
'Only Xiaohong's duck is yellow.'

(3) Zhiyou Xiaoming de fanchuan shi huangsede.
only Xiaoming DE boat is yellow
'Only Xiaoming's boat is yellow.'

(4) Zhiyou Xiaohong de shoubiao shi zisede.
only Xiaohong DE watch is purple
'Only Xiaohong's watch is purple.'

(5) Zhiyou Xiaoming de pingguo shi hongsede.
only Xiaoming DE apple is red
'Only Xiaoming's apple is red.'

(6) Zhiyou Xiaohong de maozi shi lansede.
only Xiaohong DE hat is blue
'Only Xiaohong's hat is blue.'

(7) Zhiyou Xiaoming de beizi shi lansede.
only Xiaoming DE cup is blue
'Only Xiaoming's cup is blue.'

(8) Zhiyou Xiaohong de piqiu shi hongsede.
only Xiaohong DE ball is red.
'Only Xiaohong's ball is red.'

Eight filler sentences:

(9) Zhiyou Xiaohong de shubao shi fensede.
only Xiaohong DE backpack is pink
'Only Xiaohong's backpack is pink.'

(10) Zhiyou Xiaoming de qianbi shi huangsede.
only Xiaoming DE pencil is yellow
'Only Xiaoming's pencil is yellow.'

(11) Zhiyou Xiaohong de yusan shi lansede.
only Xiaohong DE umbrella is blue
'Only Xiaohong's umbrella is blue.'

(12) Zhiyou Xiaoming de qiqiu shi huangsede.
only Xiaoming DE balloon is yellow
'Only Xiaoming's balloon is yellow.'

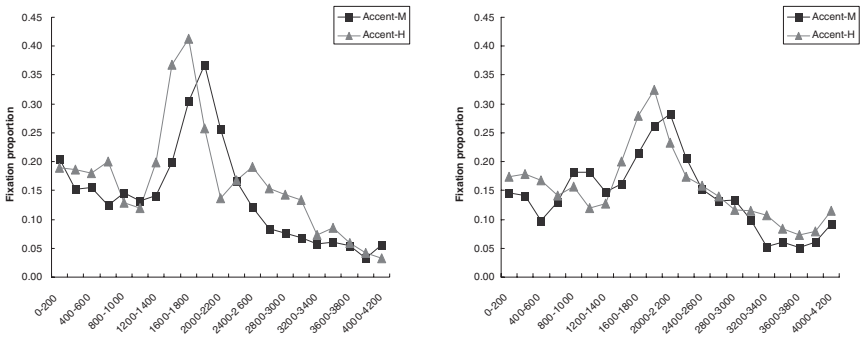
(13) Zhiyou Xiaohong de xiangpi shi fensede.
only Xiaohong DE eraser is pink
'Only Xiaohong's eraser is pink.'

(14) Zhiyou Xiaoming de xiaoxiong shi huisede.
only Xiaoming DE teddy is brown
'Only Xiaoming's teddy is brown.'

(15) Zhiyou Xiaohong de dianhua shi hongsede.
only Xiaohong DE telephone is red
'Only Xiaohong's telephone is red.'

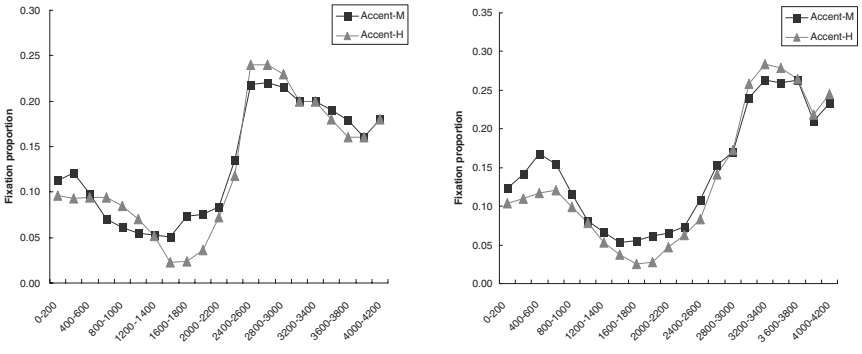
(16) Zhiyou Xiaoming de yali shi huangsede.
only Xiaoming DE pear is yellow
'Only Xiaoming's pear is yellow.'

APPENDIX B: PROPORTION OF FIXATIONS IN THE OTHER TWO CATEGORIES



Adults **Children**

Average fixation proportions over time in I (the non-contrastive object) in the two conditions by adults and children



Adults **Children**

Average fixation proportions over time in IV (the first contrastive object) in the two conditions by adults and children

APPENDIX C: TEST SENTENCES IN EXPERIMENT 2

Four test sentences:

(1) Xiaoxiongmao meiyou mai shenme jiaju.
Panda not buy what furniture

(2) Xiaojingling meiyou mai shenme shuiguo
smurf not buy what fruit

(3) Meirenyu meiyou mai shenme chongwu.
mermaid not buy what pet

(4) Jiqiren meiyou mai shenme chezi.
robot not buy what car

Filler sentences for Group A:

(5) Xiaotuzi meiyou caidao xingxing.
bunny not step star
'The bunny didn't step on the star. '

(6) Haidao meiyou nazou jinbi.
pirate not take gold-coin
'The pirates didn't take the gold coins.'

(7) Xiannü meiyou dabai dahuolong.
fairy not beat dragon
'The fairy didn't beat the dragon.'

(8) Xiaofeima meiyou zhaodao hua.
pony not find flower
'The pony didn't find the flower.'

Filler sentences for Group B:

(9) Xiaoniao chi-le shenme shiwu?
Bird eat-ASP what food
'What kind of food did the bird eat?'

(10) Xiaoma jian-le shenme baobei?
horse collect-ASP what treasure
'What kind of treasure did the horse collect?'

(11) Shei dai-le xiaomao guo he?
who take-ASP cat cross river
'Who helped the cat cross the river?'

(12) Shei dabai-le laowupo?
who beat-ASP witch
'Who beat the witch?'

APPENDIX D: CHILDREN'S RESPONSES ON THE EXAMPLE TRIAL, EXPERIMENT 2

		A rising intonation on the *wh*-phrase
Group 1	Age	Responses
Girl	4;8	(They) didn't buy wardrobes and mirrors
Boy	4;8	Pointing to the wardrobes and the mirrors
Girl	4;8	(They) didn't buy wardrobes and mirrors
Girl	4;1	Wardrobes and mirrors
Girl	4;2	Wardrobes and mirrors
Boy	4;2	(They) didn't buy wardrobes and mirrors
Boy	4;2	Wardrobes and mirrors they didn't buy
Boy	3;11	Pointing to the wardrobes and the mirrors
Girl	4;4	Wardrobes and mirrors
Boy	4;5	They didn't buy wardrobes and mirrors
Girl	3;10	(They) didn't buy wardrobes and mirrors
Girl	3;6	Wardrobe and mirrors (they) didn't buy
Girl	4;1	(They) didn't buy wardrobes and mirrors
Girl	4;2	Pointing to the wardrobes and the mirrors

Boy	4;2	Wardrobes and mirrors (they) didn't buy
Girl	4;5	Wardrobes and mirrors
Girl	4;4	Wardrobes and mirrors
Girl	4;6	(They) didn't buy wardrobes and mirrors
Boy	4;7	Wardrobes and mirrors they didn't buy

		A level intonation on the *wh*-phrase
Group 2	Age	Responses
Girl	3;7	No, (they) bought chairs
Girl	3;10	You're wrong (pointing to the chairs)
Girl	4;5	No, (they) bought chairs
Girl	4;6	No, (they) bought chairs
Boy	3;11	Wardrobes and mirrors
Boy	3;11	You're wrong, they bought chairs
Girl	4;1	No, look at the chairs
Boy	4;3	You're wrong, look at the chairs
Girl	4;4	Pointing to the wardrobes and the mirrors
Girl	4;2	No (pointing to the chairs)
Boy	4;2	No, (pointing to the chairs)
Girl	4;8	No, (they) bought chairs
Girl	4;11	You're wrong, they bought chairs
Boy	4;5	No, they bought chairs
Boy	4;2	No, they have chairs
Boy	4;4	You're wrong, they have chairs
Girl	4;5	No, they have chairs
Girl	4;6	No, (they) bought chairs
Girl	4;6	You're wrong (pointing to the chairs)

NOTE: We have translated the corresponding sentences into English. The corresponding translations are as follows.

Bu dui.
not right
'No.'

Ni shuode bu dui.
You say not right
'You're wrong.'

Meiyou mai yichu he jingzi.
not buy wardrobes and mirrors
'(They) didn't buy wardrobes and mirrors.'

Mai-le yizi.
buy-ASP chair
'(They) bought chairs.'

Kan, you yizi.
look, have chairs
'Look at the chairs.'

Tamen you yizi.
they have chairs
'They have chairs.'