Understanding prosodic focus marking in Mandarin Chinese- Data from children and adults

Hui-Ching Chen[1236], Krista Szendrői[4], Stephen Crain[35], Barbara Höhle[12]

1. International Doctorate for Experimental Approaches to Language and Brain (IDEALAB), University of Potsdam, Potsdam, Germany / University of Groningen, Groningen, Netherlands / University of Trento, Trento, Italy / University of Newcastle, Newcastle upon Tyne, United Kingdom / Macquarie University, Sydney, NSW, Australia
2. University of Potsdam, Department of Linguistics, Potsdam, Germany
3. Australian Research Council Centre of Excellence in Cognition and Its Disorders, Macquarie University, Sydney, Australia
4. University College London, Research Department of Linguistics, Division of Psychology and Language Sciences, Faculty of Brain Sciences, London, United Kingdom
5. Department of Linguistics, Macquarie University, Sydney, Australia
6. Contact information:
    - Universität Potsdam

      Department Linguistik

      Haus 14

      Karl-Liebknecht- Str. 24-25

      14476, Potsdam, Germany
    - Phone number: +49-03319772201
    - hui-ching.chen@hdr.mq.edu.au

The authors declare that there is no conflict of interest regarding the publication of this article.

How children and adults understand prosodic focus marking in Mandarin Chinese

Introduction

Following Rooth's definition, focus is used to indicate the presence of alternatives to the current element in focus in a discourse (Rooth, 1992). This function of focus can be demonstrated with question-answer pairs such as (1).

(1) Q: Who cooked this delicious soup?

A: Kathy(F)[1] cooked it (not Peter or Susan).

(2) Q: What did Kathy cook?

A: She cooked the delicious soup(F) (not the pizza or the fish)

In the felicitous answers to the questions in (1) and (2), the constituent in focus provides the answer to the question. Even though the existence of focus seems to be a shared property across languages, there is variation in the means that are used to mark focus in different languages. In languages with relatively strict word order, such as English, prosodic cues are probably the most important cue to focus marking, though syntactic constructions such as cleft sentences are used as well. In languages that allow a more flexible word order (e.g. Italian, Spanish, German) focused elements may be moved to privileged sentential positions. Nevertheless, this typically goes hand in hand with prosodic highlighting of the focused constituent (Italian: Frascarelli, 2000; Spanish: Zubizarreta, 1998; Feldhausen & Vanrell, 2014; German: Repp & Drenhaus, 2014).

Across languages, a focused constituent is often associated with prosodic salience, where salience is achieved by using phonetic properties including pitch,

[1] F means the focus of the sentence.

duration and intensity. However, the acoustic manifestations of prosodic salience and the interplay they have with other means of focus marking are subject to cross-linguistic variation and, therefore, these cues to focus have to be learned by children as they acquire the local language.

The present study asks if Mandarin-speaking children demonstrate knowledge of how prosodic information is used in sentence comprehension to identify the focused constituent in a sentence. Previous research in this area has concentrated mostly on the sensitivity to focus by English-speaking children. The findings of these investigations have suggested that children acquiring English take a somewhat paradoxical path in the course of language acquisition. Although comprehension generally precedes production in the acquisition of cognitive skills, English-speaking children appear to use prosodic cues in sentence production much earlier than they exploit these cues in sentence comprehension (e.g., Hornby, 1971; Wells et al., 2004). The paradoxical findings have not been replicated in studies of children acquiring other languages; however, recent research with French-, German- and English-speaking children, reported by Szendrői et al. (2017) has challenged the conclusion that production precedes comprehension in the acquisition of prosodic focus. The Szendrői et al. (2017) study found that children as young as 3-years-old were sensitive to prosodic salience as a means for determining the identity of a focused subject phrase. Furthermore, the study showed that speakers of English relied more heavily on prosodic information than speakers of French, with no indication of an interaction by age. In French, it turned out, the use of cleft structures is a more frequent cue to subject focus, as compared to placing a pitch accent on the subject (Hamlaoui, 2008; Lambrecht, 1994). Szendrői et al. (2017) concluded that their results reveal

cross-linguistic differences in the use of focus markers in comprehension, and that language-specific markers of focus are acquired early.

As far as we are aware, the first study investigating the comprehension of prosodically-marked focus in Mandarin children was by Chen (1998). That study reported enhanced sensitivity to prosodic focus marking by Mandarin-speaking children, as compared to adults. This difference between children and adults may be related to the fact that Mandarin is a tone language. In tone languages changes in pitch (or pitch accent) are critical cues for word recognition. Indeed, there are differences in the use of pitch as an acoustic cue for lexical tone and pitch as an intonation cue. While the domain for lexical tone is the syllable, the domain of intonation is larger than the single syllable and pitch in intonation is typically associated by other acoustic cues like intensity and duration (at least in intonation languages). However, the mulitple use of pitch in Mandarin may lead Mandarin-speaking children to attend to the prosodic properties of constituents more than adults do, and more, it could be that Mandarin-speaking children attend to pitch information more than children acquiring non-tone languages.

The present study attempts, therefore, to investigate how Mandarin-speaking children and adults exploit prosodic focus marking using materials and procedures that are similar to those used by Szendrői and colleagues (2017), where children acquiring different languages had demonstrated adult-like performance.

Mandarin Chinese is a Topic-prominent language, and a tone language. Mandarin has relatively flexible word order, for example, it permits the object phrase to be in sentence-initial (Topic) position. Given this more flexible word order and the multiple functions that pitch information have in Mandarin, one

could assume that word order makes a larger contribution than prosodic information in identifying the focused constituent. Several researchers have reached precisely this conclusion (Feng, 2003; Shyu, 2012; Xu, 2004). For example, Xu (2004) proposed that Mandarin has a default focus position, which is the final position in the most embedded clause. If a stressed constituent appears in this position, it is doubly marked for focus (by syntactic and prosodic means) and typically constitutes contrastive focus.

However, it is worth noting that focus can also be marked solely by prosody in Mandarin. For example, in canonical SVO-sentences in which the subject is the focus element, the subject is prosodically highlighted. As observed by both Shyu (2012) and Xu (2004), however, focus is sometimes dissociated from prosodic salience in Mandarin; that is, it is possible to express focus without any phonological manifestation. This separates Mandarin from so called focus-stress languages which typically show a reliable association between focus and prosodic prominence. The separation of focus information and prosodic marking may be due to the fact that Mandarin is a Topic-prominent language (Li and Thompson 1981, p. 15). In contrast to Subject-prominent languages such as English, the sentence initial position in Topic-prominent languages typically contains the Topic; what the sentence is about. In general, pitch accent plays only a minor role in encoding topics in Asian languages (Féry and Krifka, 2008).

Empirical studies confirm that Mandarin speakers use prosodic information in their realization of focus. In contrast to West-Germanic languages, where pitch accent is typically used to mark focus, focus is not only signalled by an extended pitch accent in Mandarin, but also by longer duration and higher intensity (e.g. Chen & Braun, 2006; Chen & Gussenhoven, 2008; Ouyang & Kaiser,

2014; Xu, 1999). Yang and Chen (2014) showed that even 4-year-old Mandarin-speaking children used pitch, duration, and intensity to mark focus. This observation is in line with the finding that even very young children acquiring focus-stress languages produce prosodic focus markers.

As compared to the use of prosody in sentence production, less is known about how Mandarin speakers exploit prosodic information in identifying the focus constituent in sentence comprehension. A study by Chen (1998) used a sentence-picture verification task and a sentence-correction task to assess the use of different syntactic structures and prosodic information by Mandarin-speaking children and adults. This study examined the sensitivity of the participants in determining given vs. new information in the test sentences. Following Hornby's study (1971) with English-speaking children, Chen (1998) examined four types of transitive sentences: canonical SVO active sentences, passive sentences, cleft sentences and pseudo-cleft sentences. Each sentence type was either presented with neutral prosody or with emphatic stress on either the subject or the object of the sentence. For the purposes of the present study, the results from the active SVO sentences are the most relevant. The hypothesis in this study was that the final constituent (the object) would be interpreted as new information in the neutral prosody condition. The crucial question was whether putting an emphatic stress specifically on the sentence subject would change this strategy. It turned out, however, that adult Mandarin-speakers were not influenced by the addition of emphatic stress. Instead, the adult participants consistently interpreted the sentence object as new in both the neutral prosody condition and in the conditions that added emphatic stress. However, the 5- to 13-year-old Mandarin-speaking child participants tended to interpret the subject as new when that constituent

carried an emphatic stress. Based on these results, Chen (1998) concluded that Mandarin-speaking children rely more heavily on prosodic information than on positional information in determining given versus new information, whereas Mandarin-speaking adults considered word order to be the major cue for given vs. new information.

A study by Zhou et al. (2012) investigated the sensitivity of 5-year-old Mandarin-speaking children, and a control group of adults, to the use of prosodic information in on-line sentence comprehension. The aim of the study was to determine which sentence constituent was associated with the Mandarin counterpart to the English focus adverb *only* (viz., Mandarin *zhiyou*). In the test sentences, prosodic emphasis was placed on the head noun of the subject phrase, or on a modifier of the subject phrase. In English, these different associations can be illustrated by the following sentences (where capital letters indicate prosodic stress): *Only John's APPLE is red* versus *Only JOHN'S apple is red*. The task was to judge whether or not the test sentences matched a visual display. The visual displays contained depictions of objects that were natural alternatives to the modifier of the subject phrase or to the head noun of the subject phrase. An eye-tracker was used to measure participants' gaze durations to the alternative objects. The hypothesis was that when the modifier of the subject phrase received prosodic stress the participants would look longer at the objects that were alternatives to the modifier of the subject phrase comparing to when the head noun of the subject phrase received prosodic stress. On the other hand, the participants would look longer at the objects that were alternatives to the head noun of the subject phrase when it was stressed comparing to when the modifier of the subject phrase was stressed. The results confirmed this expectation both for

4-to-5-year old child participants, as well as for the adult participants. Nevertheless, the judgment data revealed significantly different patterns of behavioral responses by the two participant groups.

The judgement pattern by the adult participants was exactly as predicted by the hypothesis and driven by the stress information. However, the behavioral responses by the child participants did not appear to be influenced by the different stress patterns in the test sentences. The child participants consistently favored the interpretation that was associated with stress on the modifier of the subject phrase. That is, children responded to both kinds of test sentences in the same way, although their eye-movements patterns suggested sensitivity to the stress information. The child participants associated the Mandarin focus adverb *zhiyou* with the modifier of the subject phrase, despite having looked at the alternatives to the head noun when it received stress. It appears, then, that Mandarin-speaking children are sensitive to different stress patterns, but cannot use this sensitivity as a guide to sentence interpretation.

The findings of previous studies therefore draw an inconsistent picture of the sensitivity by Mandarin-speaking children to prosodic cues in identify information structure in sentences. The findings of the study by Chen (1998) suggest that Mandarin children show a stronger reliance on prosodic information than adults do when identifying new information. In contrast, the findings of the study by Zhou et al. (2012) suggest that 4- to 5-year-old Mandarin-speaking children, like adults, are sensitive to prosodic information, but this sensitivity is limited to on-line processing, and is not revealed in children's behavioral responses.

The current study

The goal of the present study is to gain a clearer picture about the use of prosodic information by Mandarin-speaking children and adults to identify the focus constituent of sentences. The design of the study was based on the experimental protocols used by Szendrői et al. (2017). As in that study, the participants in the present study were presented with simple subject-verb-object sentences in which prosodic emphasis was placed on the subject noun phrase or on the object noun phrase. They were presented as descriptions of pictures that either matched the sentence, or did not match it. The participants' task was to judge whether or not the sentence matched the picture and, if the sentence did not match the picture, to correct the sentence. On each trial, the information contained in the picture permitted the participants to correct either the subject or the object. If Mandarin speakers are sensitive to pitch accent as a marker of the element in focus, the participants were expected to correct the subject in sentences where the subject received the pitch accent, and were expected to correct the object in sentences where the object received the pitch accent. Because Mandarin is a Topic-prominent language, there is a strong preference for Mandarin speakers to place the focused element in sentence final position (Xu, 2004). Based on this understanding, we anticipated that participants would exhibit an overall preference to assign focus to the object phrase, rather than to the subject phrase. If so, there would be a higher proportion of corrections of the object phrase, even when the pitch accent was placed on the subject phrase (the Topic). This last prediction could be expected to hold to a greater extent for the adult participants than for the child participants, if Chen (1998) is correct in inferring that Mandarin-speaking children are more reliant than adults are on the use of prosodic information to identify the focused constituent of a sentence. This study intends

to contribute to a better understanding of prosodic focus marking and its development in tonal and Topic-prominent languages in which this topic has not been thoroughly investigated. Further, using the same experimental setup that has already been used across some other languages will allow us to set the findings of Mandarin learning children in relation to those from children learning non-tonal languages. Thus, it sheds more lights on the language-specific and language independent aspects of focus development.

Method

Participants

All participants were native speakers of Mandarin Chinese living in Kaohsiung, Taiwan. There were six groups of participants. First, 20 university students from National Chong-Shan University participated in an experiment that placed prosodic emphasis on the subject phrase of the test sentences. This will be called the subject-accented condition. Another 20 university students from National Chong-Shan University participated in an experiment that placed prosodic emphasis on the object phrase of the test sentences. This will be called the object-accented condition[2]. Comparable to the Szendrői et al. study (2017) and the Chen study (1998), we collected data from four groups of children. A group of 20 3- to 4-year-old children was tested in the subject-accented condition (M = 3;6, age range: 3;1-4;11). A group of 21 3- to 4-year-old children was tested in the

[2]All university students started to learn English as a foreign language at age twelve. They were students from the Department of Sinology and had an infrequent use of English by the time of testing.

object-accented condition (*M* =3;8, age range: 3;3-4.10). A group of 23 5-year-old children was tested in the subject-accented condition (*M* = 5;5, age range: 5-5;11). Finally, a group of 21 5-year-old children was tested in the object-accented condition (*M* = 5;6, age range: 5-5;11). The groups with 3- to 4-year-old children are called the younger children group later.

Materials and design

Sentence-picture verification task. The experiment consisted of 14 trials. These were comprised of two practice trials, four target trials, four control trials and four filler trials.

To investigate whether Mandarin-speaking children are sensitive to prosodic cues as a vehicle for identifying the focus of a sentence, we adapted the sentence-picture verification task used by Szendrői et al. (2017) for Mandarin-speaking participants. On each trial, the participant saw a coloured picture presented on a 15-inch laptop screen (see Figure 1). Each picture displayed three different animals with an object close to it. At the same time, the participant heard a test sentence, produced by a trained female experimenter.

Each participant was either tested in the subject-accented condition (see Examples 1 and 3) or in the object-accented condition (see Examples 2 and 4). The participants were randomly assigned to one of these conditions. The test sentences in the four target trials were false descriptions of the pictures in both conditions, if the participant used pitch accent to guide the assignment of the focus constituent (see Examples 1 and 2). Only these false sentences could reveal a correct focus assignment as the participants had to correct false statements. In contrast, the control sentences were true descriptions of the corresponding

pictures to balance the number of true and false statements (see Examples 3 and 4), for the participants who assigned focus by exploiting the prosodic information.

Fig. 1 Example of the visual stimuli

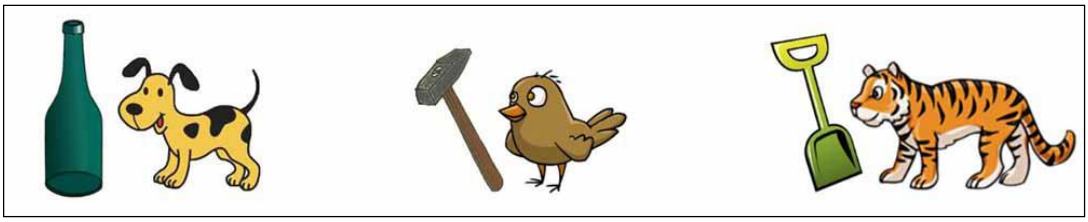


Table 1 Examples of the test sentences used as target and control trials[3]

Subject-accented sentence	Object-accented sentence
Example1: Target Trial	Example2: Target Trial
XIAONIAO$_F$ you shueping, shi ma?	xiaoniao you SHUEPING$_F$, shi ma?
BIRDY$_F$ has bottle, Aux Q?	birdy has BOTTLE$_F$, Aux Q?
'The BIRDY$_F$ has the bottle, is that right?'	'The birdy has the BOTTLE$_F$, is that right?'
Focus congruent response:	Focus congruent response:
'No, the doggie has the bottle.'	'No, the birdy has the hammer.'
In Mandarin:	In Mandarin:
Bushi, xiaogou you shueping.	Bushi, xiaoniao you chuizi.

[3] As before, in examples 1, 2, 3 and 4, F means focus of the sentence. Furthermore, capital letters mark an accented word.

Example 3: Control Trial	Example 4: Control Trial
XIAONIAO$_F$ you chuizi, shi ma?	xiaoniao you CHUIZI$_F$, hi ma?
BIRDY$_F$ has hammer, Aux Q?	birdy has HAMMER$_F$, Aux Q?
'The BIRDY$_F$ has the hammer, is that right?'	'The birdy has the HAMMER$_F$, is that right?'
Response: 'Yes'	Response: 'Yes'

The pictures used in the filler trials were similar to the pictures corresponding to the target and control trials. However, the filler sentences mentioned a property that was shared among all the animals or objects that were depicted in the pictures, such that the participant was required to check all the animals and objects in order to make an accurate judgment. For example, the visual scene depicted in Figure 1 was shown in conjunction with the sentence: *All animals are green, is that right?* In this case the expected answer would be *No, the doggie is yellow, the birdy is brown and the tiger is orange*. There were two true and two false statements for the filler trials. The three types of trials were presented in a pseudo-randomized order so that no more than one trial of one type followed each other. The participants were asked to decide whether the sentence was a true description of the picture. If not, they were required to correct the sentence by using a SVO structure, such as, *No, the birdy has the hammer*. or *No, the doggy has the bottle*.

All pictures depicted three pairs of animals and objects. Within each animal-object pair, the animals were always positioned on the right-hand side of the picture and the objects were positioned on the left-hand side. In the subject-accented condition (see Example 1) the animal which corresponded to the

constituent in focus in the test sentence was always in the middle of the three animal-object pairs. In contrast, in the object-accented condition (see Example 2), the object which corresponded to the constituent in focus in the test sentence appeared twice on the left-hand side of the display and twice on the right-hand side. Each animal and each object was only used in one target trial. For the control and filler trials, the animals (but not the objects) appeared in more than one picture, but never with the same combination of animals or objects. Some of the animals and the objects used in the Szendrői et al. (2017) study were replaced, so that they would be familiar to young children acquiring Mandarin (e.g. *hedgehog, lollypop*). Parents were asked to fill in a parental questionnaire after the experiment and this confirmed that the animals and the objects used in the experiment were known to their children. This task is not only apt to test younger participants, in our case, three years old children but also to grant focus reading (Szendrői et al., 2017).

Procedures

All participants were tested individually in a quiet room either at the university or in the kindergarten. Each participant was seated in front of a laptop computer, on which the visual stimuli were presented. The experimenter was located behind the screen, such that only the participant could see the pictures. The experimenter told the participant that she (the experimenter) would try to remember what was in the pictures but that she wasn't sure how successful she would be. Therefore, the participant was asked to judge whether the sentence presented by the experiment was an accurate description of each picture. The participant was asked to correct sentences that were not accurate descriptions of the pictures. The stimulus sentences were produced by the experimenter (and not pre-recorded) to

make the situation as natural as possible. Each response from the participant was audio-recorded, manually transcribed and coded. To familiarize the participants with the sentence-picture verification task, the experimenter started with two practice trials which required one Yes and one No response. The study was approved by the ethics committee of the University of Potsdam.

Coding responses

Corrections of the subject phrase in the subject-accented condition and corrections of the object phrase in the object-accented condition were considered to be focus-congruent responses, and were scored as correct. In contrast, corrections of the object phrase in the subject-accented condition and corrections of the subject phrase in the object-accented condition were classified as focus-incongruent responses, and were scored as incorrect. In four cases, participants corrected both phrases (e.g. *No, the doggie has the bottle and the birdy has the hammer*). In such cases, the response was coded as invalid, and excluded from statistical analysis. Although participants generally produced full SVO sentences in making their corrections, several elliptical responses were also produced, consisting solely of a noun phrase (e.g. *the dog, the bottle*). Elliptical responses were especially characteristic of the groups of younger children. These responses were counted as valid, however, since it was clear which phrase was being corrected, so these responses were scored in the same way as full sentences, according to their (in)congruence with the constituent that was prosodically accented. That is, if a participant answered *the dog* in the subject-accented condition, or *the bottle* in the object-accented condition, then the answer was scored as correct.

Results

All responses in the control and filler trials were correct, so no participants' results were excluded from statistical analysis. For inferential statistics, we analyzed the data using a linear mixed-effects model (Baayen et al., 2008) in R (R Core Team, 2016) with the lme4 package (Bates et al., 2015) in the R-Studio environment (Version 1.0.136). Our aim was to assess the impact of age and sentence condition, and their interaction, on the number of focus congruent responses produced by participants in the test trials. In addition to these fixed effects, the model contained two random effects, which can be viewed as accounting for variation in responses by individual participants and to individual items (random intercepts). We examined whether or not the proportion of focus-congruent responses differed significantly between groups in the subject-accented condition, and whether there were significant differences in the proportions of focus-congruent responses between the subject-accented condition and the object-accented condition[4].

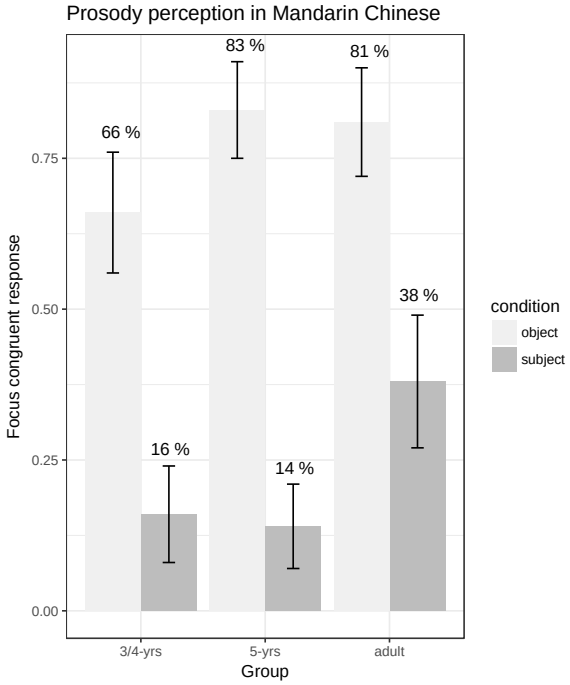
The data in the subject-accented condition of the younger group was used as the baseline in the model. Overall focus congruent responses were significantly lower in the subject-accented condition than in the object-accented condition ($Group_{3\&4}$: b=-3.78, SE=0.85, Z=-4.432, $p<0.001$). Further, the performance of the younger children did not differ significantly from the performance of the 5-year-olds in the subject-accented condition (b=-1.23, SE=0.75, Z=-1.642, $p>0.1$), but there was a significant difference between the younger children and the adults (b=1.01, SE=0.50, Z=-2.025, $p<0.05$). Furthermore, there was a significant interaction (Group*Condition) between the group of younger children and the

[4]Response~group*condition+(1|id)+(1|item)

group of the older children (b=2.49, SE= 1.13, Z=2.198, p<0.05), but not between the group of younger children and the adults (b= - 0.44, SE= 0.89, Z=-0.498, p>0.61). These results are summarized in Figure 2.

Across all age groups, the participants in the object-accented condition produced significantly more focus-congruent responses than participants in the subject-accented condition. In addition, the adult participants performed more accurately than both groups of child participants in the subject-accented condition. The interaction between the 3- to 4-year-old group and the 5-year-old group reflects the larger difference in performance between the object- and subject-accented conditions for the older children, as compared to the younger children. This was mainly due to the lower performance of the 3- to 4-year-old group in the object-accented condition. The absence of an interaction between the 3- to 4-year-olds and the adults indicates that the differences in performance between the two conditions are similar for these groups due to the higher performance in the subject-accented condition of the adult group.

Fig. 2 Percentage of focus congruent responses across sentence conditions and participant groups. Each error bar showed the 95% confidence intervals for the mean.



Discussion

The present study investigated the sensitivity of Mandarin-speaking adults and 3-, 4- and 5-year-old Mandarin-speaking children to pitch accent as a vehicle for marking the focused constituent in a sentence. To address the issue, sentence-picture verification task in which the participants were asked to correct sentences that did not match the associated pictures was conducted. The pictures were constructed to permit corrections of either the subject phrase of the sentence or the object phrase. The hypothesis was that – if participants use stress for focus assignment – they would produce a higher rate of corrections for the subject phrase if it was accented, and a higher rate of corrections for the object phrase when it was accented. However, if stress did not affect focus assignment we expected that the number of corrections for the object phrase would be

significantly higher than the corrections for the subject phrase. This prediction was based on the fact that, as a Topic-prominent language, Mandarin speakers may analyze the constituent in sentence-initial position as the Topic (i.e., what the sentence is about), and analyze the constituent in sentence-final as the focus (a comment about the Topic).

All of the groups exhibited a significantly higher proportion of focus-congruent responses in the object-accented condition, as compared to the subject-accented condition which is in line with our second prediction. The number of focus-congruent responses in the subject-accented condition was higher for adults than for either group of children. The first main finding of the study indicates that Mandarin speakers make little use of prosodic information to identify the focus of the sentence and, instead, rely on word order to determine the constituent that is in focus. In the present experiment, the focused constituent is predominately taken by the participants to be the last constituent of the sentence – the object phrase. The second main finding is that the adult participants were more flexible than children in the assignment of focus, as adults exhibited more focus-congruent responses than children in the subject-accented condition, where the focus accent was on the subject phrase of the test sentences.

These findings are in line with the proposal that the use of prosodic focus marking is quite restricted in Mandarin Chinese and that word order is more widely used to indicate the (topic/comment) information structure of sentences. It should be noted, however, that prosodic prominence is used to signal non-default focus on the subject phrase of a sentence (Shih 1988; Xu 1999). Our results demonstrate that speakers of Mandarin – children as well as adults – rely more on word order than on prosodic information for the assignment of focus and that

children are even less facile than adults are in the use of prosodic cues to mark the constituent that is in focus.

Overall the findings are only partly compatible with those of Chen (1998). Chen did find that adults relied more heavily on word order than on prosodic information in the active sentence condition, which is the condition that is most similar to the test sentences used in the present study. Even when emphatic stress was placed on the subject phrase of the active sentences in the Chen study, the adult participants corrected the sentence final object phrase roughly 70% of the time. We observed a similar pattern in the overall higher correction rate (62%) for the object phrase in the subject-accented condition.

Despite the similar patterns of responses by the adult participants in the two studies, the data from the child participants contrasted sharply. The child participants in the Chen study corrected the subject phrase roughly 65% of the time when emphasis was placed on the subject phrase. By contrast, the group of 5-year-olds in the present study corrected the subject phrase in the subject-accented condition only 14% of the time. Thus, there was no indication from our data that children make more use of prosodic information than adults do, as Chen (1998) had previously concluded.

We have no unequivocal explanation why the results of the two studies are so discrepant, since both the testing procedures and the age of the child participants were quite similar, at least when we limit attention to the group of 5-year-old child participants in our study. It should be noted, however, that Chen's experiment involved a much higher number of different experimental conditions. Our study included only semantically irreversible active SVO sentences with stress placed either on the subject phrase or on the object phrase. By contrast, the Chen

study included 96 semantically reversible and irreversible actives, passives, as well as cleft and pseudocleft sentences. In addition, the test sentences in the Chen study were presented in three prosodic conditions. The high number of sentences, the greater structural diversity, and the additional prosodic condition could have made the task in the Chen study more challenging than that of the present study. It is feasible that the child participants responded to the challenge with increased attention to the acoustically more salient constituent of the test sentences. This in turn may have led them to correct this constituent irrespective of the information structure of the sentences.

The findings from the present study also only partly align with those from Zhou et al. (2012). The Zhou et al. (2012) study found that, although the gaze patterns of 4-year-old children and adults were similar, the patterns of behavior by children and adults differed. Only the adult participants, but not the child participants, attended to the placement of focal stress as the basis of their behavioral responses. This finding is comparable with the finding in the present study that adults showed a stronger sensitivity to focal stress than children. However, the Zhou et al. (2012) study also revealed a discrepancy in the pattern of responses by the two age groups dependent on the experimental method. Children and adults performed differently in the behavioral test, but they displayed similar patterns of behavior in the eye-tracking task. The observation that methodological aspects of experiments that require pragmatic skills have a large impact on children's performance has been found in several other studies (Höhle et al., 2016; Berger & Höhle, 2012; Papafragou & Musolino, 2003).

The present study adopted similar experimental materials and the same procedures as the study by Szendrői et al. (2017) who tested children acquiring

German, French, and English. This permits us to include our findings from Mandarin-speaking children in a cross-linguistic comparison of child language. The findings from children in the present study augment the previous data in several ways. Most importantly, the findings from the present study further support the general picture obtained in the Szendrői et al. study according to which children's performance in focus interpretation resembles the language specific pattern from early in the course of language development. No interactions with age or language were found in the Szendrői et al. study, suggesting that the participants used prosodic information as adults did to identify the focus constituent in a sentence in the same way across languages and across the ages that were tested. In the present study with Mandarin, by contrast, we found an interaction by age. This interaction revealed a larger difference in performance between the subject-accented and object-accented conditions for the group of 5-year-olds, as compared to both the groups of younger children, and the group of adults. However, two opposing trends caused the interaction: the groups of 3- and 4-year-olds had slightly fewer focus congruent responses in the object-accented condition, whereas the group of adults exhibited a higher number of focus congruent responses in the subject-accented condition. Notably, the differences in the two conditions across age groups was smaller when the group of adults was compared to the group of 5-year-olds, but this was due to different reasons. The conclusion is that children's early adherence to prosodic prominence as a focus marker is not due to a universal bias for children to attend to acoustically highlighted information. Rather, even young children's strategies for identifying the focus of sentences reflects language specific properties. Young children acquiring French, German and English recognize that they are acquiring stress-

focus languages, and young Mandarin-speaking children recognize that focus is largely determined by structural position in Asian languages.

Another commonality between the two studies is the general tendency for children to assign focus to the sentence object. This tendency appears to be independent of the specific strategies adopted across languages, and is not negated when the subject phrase carries focal stress. A cross-linguistic comparison of the patterns of responses by adults in the subject-accented condition suggests that the language specific balance in the use of prosodic versus word order information predicts their performance. That is, adult speakers of English - a language with highly rigid word order - produced 73% focus-congruent responses, adult speakers of German 59% and adult speakers of French only 32.5% focus-congruent responses. Thus the adult speakers of Mandarin and the speakers of French were most similar in their percentages of focus-congruent responses. Interestingly, both French and Mandarin use specific syntactic structures such as the cleft construction, to mark subject focus (Hamlaoui, 2008; Lambrecht, 1994) such that prosodic focus marking may not be so salient for speakers of these two languages.

The general preference for object focus across all languages support the conclusion that subject focus is a marked structure. For stress-focus languages, Reinhart (2004) has proposed that this asymmetry has prosodic reasons as the neutral position for prosodic prominence is the mostly embedded constituent and thus the final position in SVO sentences. The subject phrase can only get prosodic prominence by stress shift. Thus, sentences with subject focus involve stress shift, and are therefore likely to be computationally more costly than ones with object focus. If this relationship between markedness and stress shift also holds for

Mandarin, this could explain the observed difference between Mandarin-speaking children and adults. In this way, children's greater limitations in computational resources cause them to rely more on less costly word-order strategies for focus assignment, whereas adults' extended computational resources allow them to include prosodic information into sentence interpretation.

To conclude, the present study adds to our cross-linguistic understanding of how the assignment of focus develops in the semantic representations of children, extending previous research on children who are acquiring a Topic-prominent and a tonal language. As demonstrated, the findings of the present study are inconsistent with previous findings that children acquiring Mandarin were more sensitive than adult speakers of Mandarin to prosodic cues as marker for information structure. Instead, we have demonstrated that children relied more on word order information than on prosodic information while understanding focus in sentences. This suggests that children's sensitivity to focal stress found in other languages is not an extra-grammatical, purely acoustically driven vehicle that enables children to attend to salient aspects of the signal. Moving forward, more research is needed to support these conclusions, including studies that use comparable experimental designs, and studies that include children acquiring languages that adopt different strategies in determining information structure.

Acknowledgements

We would like to acknowledge the German Exchange Academic Service (DAAD) for Chen, the DFG SFB632, and the support of the ESF EURO-EXPRAG Research Network Program for Höhle and Szendrői. Additionally, this work was also

supported by the Erasmus Mundus Joint Doctoral Programme of the European Union (IDEALAB), 2014-0685/001-001-EMJD (Framework Partnership Agreement 2012-2015). We also acknowledge the ARC Centre of Excellence in Cognition and its Disorders and as well as the Babylab in Potsdam. Finally, we would like to thank all the participants, parents and the teachers involved in the study.

References

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and S4. *R package version, 1*(7), 1-23.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390-412.

Berger, F., & Höhle, B. (2012). Restrictions on addition: children's interpretation of the focus particles auch 'also'and nur 'only'in German. *Journal of Child Language, 39*(2), 383-410.

Chen, S. H. E. (1998). Surface cues and the development of given/new interpretation. *Applied Psycholinguistics, 19*(4), 553-582.

Chen, Y., & Braun, B. (2006). Prosodic realization of information structure categories in standard Chinese. *Speech Prosody* (p. 54).

Chen, Y., & Gussenhoven, C. (2008). Emphasis and tonal implementation in Standard Chinese. *Journal of Phonetics, 36*(4), 724-746.

Feldhausen, I., & Vanrell, M. (2014). Prosody, focus and word order in Catalan and Spanish. An Optimality Theoretic approach. Paper presented at the *Proceedings of the 10th International Seminar on Speech Production*.

Feng, S. (2003). Prosodically constrained postverbal PPs in Mandarin Chinese. *Linguistics, 41*(6; ISSU 388), 1085-1122.

Féry, Caroline & Manfred Krifka. (2008). Information structure. Notional distinctions, ways of expression. In Piet van Sterkenburg (ed.), *Unity and Diversity of Languages*, Amsterdam: John Benjamins, 2008, 123-136.

Frascarelli, M. (2000). The syntax-phonology interface in focus and topic constructions in Italian (Vol. 50): *Springer Science & Business Media*.

Hamlaoui, F. (2008). Focus, contrast and the syntax-phonology interface: The case of French cleft sentences. Paper presented at the *Proceedings of the 18th International Congress of Linguists*.

Höhle, B., Berger, F., & Sauermann, A. (2016). Information Structure in First Language Acquisition *The Oxford Handbook of Information Structure*.

Hornby, P. A. (1971). Surface structure and the topic-comment distinction: A developmental study. *Child Development*, 1975-1988.

Lambrecht, K. (1994). Information structure and sentence form: Topic, focus, and the mental representations of discourse referents (Vol. 71): *Cambridge University Press.*

Li, Charles & Thompson, Sandra A. (1981). Mandarin Chinese: A Functional Reference Grammar. Berkeley, Los Angeles and London: *University of California Press*.

Ouyang, I. C., & Kaiser, E. (2014). Prosody and information structure in a tone language: an investigation of Mandarin Chinese. *Language Cognition and Neuroscience, 30*(1-2), 57-72.

Papafragou, A., & Musolino, J. (2003). Scalar implicatures: experiments at the semantics–pragmatics interface. *Cognition, 86*(3), 253-282.

Reinhart, T. (2004). The processing cost of reference set computation: Acquisition of stress shift and focus. *Language Acquisition, 12*(2), 109-155.

Repp, S., & Drenhaus, H. (2014). Intonation influences processing and recall of left-dislocation sentences by indicating topic vs. focus status of dislocated referent. *Language, Cognition and Neuroscience, 30*(3), 324-346.

Rooth, M. (1992). A theory of focus interpretation. *Natural Language Semantics, 1*(1), 75-116

Shih, C. (1988). Tone and intonation in Mandarin. *Working Papers of the Cornell Phonetics Laboratory, 3*, 83-109.

Shyu, S. I. (2012). Topic and focus. *The Handbook of Chinese Linguistics*, 100-125.

Szendrői, K., Bernard, C., Berger, F., Gervain, J., & Höhle, B. (2017). Acquisition of prosodic focus marking by English, French, and German three-, four-, five- and six-year-olds. *Journal of Child Language*, 1-23.

Wells, B., Peppé, S., & Goulandris, N. (2004). Intonation development from five to thirteen. *Journal of Child Language, 31*(4), 749-778.

Xu, L. J. (2004). Manifestation of informational focus. *Lingua, 114*(3), 277-299.

Xu, Y. (1999). Effects of tone and focus on the formation and alignment of F 0 contours. *Journal of Phonetics, 27*(1), 55-105.

Yang, A., & Chen, A. (2014). Prosodic focus marking in Chinese four-and eight-year-olds. Paper presented at the *Speech Prosody*.

Zhou, P., Su, Y., Crain, S., Gao, L. Q., & Zhan, L. K. (2012). Children's use of phonological information in ambiguity resolution: a view from Mandarin Chinese. *Journal of Child Language, 39*(4), 687-730.

Zubizarreta, M. L. (1998). Prosody, focus, and word order: *MIT Press*.

Fig. 1 Example of the visual stimuli

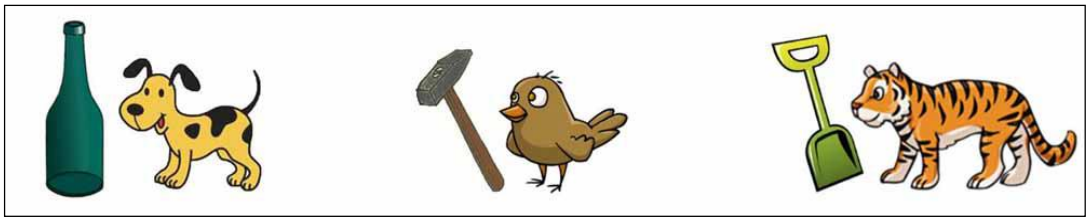


Fig. 2 Percentage of focus congruent responses across sentence conditions and participant groups. Each error bar showed the 95% confidence intervals for the mean.

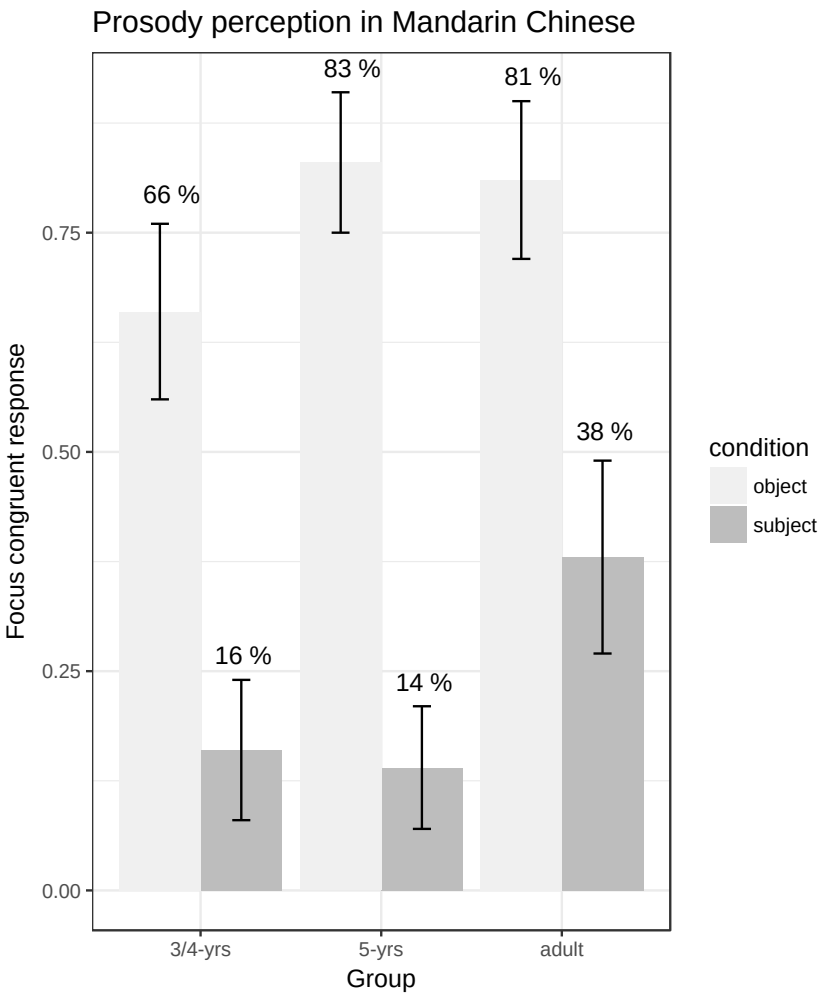


Table 1 Examples of the test sentences used as target and control trials

Subject-accented sentence	Object-accented sentence
Example1: Target Trial XIAONIAO$_F$ you shueping, shi ma? BIRDY$_F$ has bottle, Aux Q? 'The BIRDY$_F$ has the bottle, is that right?' Focus congruent response: 'No, the doggie has the bottle.' In Mandarin: Bushi, xiaogou you shueping.	Example2: Target Trial xiaoniao you SHUEPING$_F$, shi ma? birdy has BOTTLE$_F$, Aux Q? 'The birdy has the BOTTLE$_F$, is that right?' Focus congruent response: 'No, the birdy has the hammer.' In Mandarin: Bushi, xiaoniao you chuizi.
Example 3: Control Trial XIAONIAO$_F$ you chuizi, shi ma? BIRDY$_F$ has hammer, Aux Q? 'The BIRDY$_F$ has the hammer, is that right?' Response: 'Yes'	Example 4: Control Trial xiaoniao you CHUIZI$_F$, hi ma? birdy has HAMMER$_F$, Aux Q? 'The birdy has the HAMMER$_F$, is that right?' Response: 'Yes'