ACQUISITION OF FOCUS - IN A CROSS-LINGUISTIC PERSPECTIVE

By

Hui-Ching Chen

A dissertation submitted to the Faculty of Human Sciences

University of Potsdam

In Fulfilment of the Requirements for the Degree of

Doctor of Philosophy

2022

Supervised by Professor Barbara Höhle and Professor Stephen Crain

Acknowledgement

Every journey starts with excitement and ends of gratitude, and I would like to express my sincere appreciation to my supervisors, Prof. Barbara Höhle and Prof. Stephen Crain, for their continuous support of my PhD journey. Without their support, this PhD journey would not be able to come to an end, especially during this difficult time. They have been expanding my view of being a researcher as well as generously sharing their knowledge with me. I would continuously apply what I have learned from them and hopefully I will share this valuable experience and knowledge with others in the future. More words cannot truly express my gratitude and I will keep them in mind and help others when there is any other opportunity. Secondly, I would like to thank Prof. Lyndsey Nickels, Prof. David Howard, Prof. Gabriele Miceli, Prof. Roelien Bastiaanse for their substantial feedbacks and support during my PhD. I would not forget every moment that I could spend with you during the summer school's penal and winter school's penal. Additionally, I would also like to thank the German Exchange Academic Service (DAAD), the DFG SFB632, and the Erasmus Mundus Joint Doctoral Programme of the European Union (IDEALAB), 2014-0685/001-001-EMJD (Framework Partnership Agreement 2012–2015) for their financial support.

More, I would like to thank my Potsdam colleagues and my Macquarie colleagues, Tom, Antje, Maja, Ines, Jeannine, Silvana, Annett, NatalieBoll-Avetisyan, Ines, Anje, Antonia, Jeannine, Silvana, Annett, Elina, Robinson, Lesley, Leanne, Ana, Na, Valerie, and Lyn. Thank you for your help and every wonderful discussion during lunch break, coffee break, Christmas celebration, etc. Those moment have motivated me a lot! There is a saying in Chinese, "At home you depend on your parents, outside you rely on your friends".

Further, I would like to thank you all my dearest IDEALAB colleagues, with you, I am not along on the journal all over the world, Assunta, Nenad, Fleur, Rowena, Bernard, Katarina, Inga, Fleur, Ana, Ella, Jakolien, Serine, Alexa, Hahn and a lot of you! I remembered once I was at a conference, and I was sharing with one of the other PhD candidate students that it was fairly easy to accommodate myself at the guest university when I was doing my research there and she was saying it was very difficult for her in the beginning while she was doing hers at the guest university. It was my super IDEALAB colleagues who make everything easier for my PhD journey. The memory that we created together has also become my superpower to support me to continue my dream ahead.

I would also like to thank the labs and schools, i.e., Reading and Eye-tracking Lab at National Chang-Chi University, Brain and Language Laboratory at National Central University, Department of Education Psychology and Counseling at National Normal Unviersity, Zhengbei and ZhongShan primary school, Yumiao, Peixian, Zhonghe and Yixian kindergarten in Taiwan where I was allowed to conduct the experiments in Taiwan. Besides, I thank all the participants, the teachers, the children, and their parents who either helped to set up the experiment or participate in the studies. Thank you for your efforts and support. Additionally, I would like to thank Dr. Yow and my current colleagues at the Social and Language Cognition Lab at Singapore University of Technology and Design for their support to company me to the end of this PhD journey.

Finally, I would like to thank my parents, my sister, my grandparents, my uncles, my aunties, and my cousins for supporting me to pursue my dreams. Because of your support, I could come this far and there is saying in Mandarin, "A journey of a thousand miles begins with a single step." and with this little step and your support, I will be able to further follow my dream and hopefully I could apply the knowledge that I gain to help others, and just like the others have helped me too.

Table of Contents

Abbreviations

NP: noun phrase

PFV: perfective aspect (-le)

Q: question(ma)

SOV: subject-object-verb

SVO: subject-verb-object

Abstract

Successful communication is often explored by people throughout their life courses. To effectively transfer one's own information to others, people employ various linguistic tools, such as word order information, prosodic cues, and lexical choices. The exploration of these linguistic cues is known as the study of information structure (IS). Moreover, an important issue in the language acquisition of children is the investigation of how they acquire IS. This thesis seeks to improve our understanding of how children acquire different tools (i.e., prosodical cues, syntactical cues, and the focus particle only) of focus marking in a cross linguistic perspective.

Following Rooth (1992), focus indicates the alternative in the discourse, and can be assigned by various linguistic means. Question and answer pairs are often used for implying focus of a sentence. For example,

(1) Question: Who wrote a book about focus interpretation?

Answer: ROOTH wrote the book about focus interpretation.

Here, the focus in the answer is the subject 'Rooth'. An English speaker could either add an accent on the subject, i.e., '*ROOTH* wrote the book about focus interpretation, (but not Krifka or Gussenhoven)', or employ cleft sentence structure, i.e., *'it's* ROOTH who wrote the book about focus interpretation, (but not Krifka or Gussenhoven)'. Another method of assigning focus is through the focus particle, i.e., '*Only* ROOTH wrote the book about focus interpretation, (but not Krifka or Gussenhoven)*'*. As part of this research, we conducted three studies to scrutinise how tone and intonation language speakers (Mandarin and German speakers, respectively) acquire these linguistic tools.

In the first study, following Szendrői and her colleagues (2017), we performed a sentence-picture verification task to investigate whether three- to five-year-old Mandarin-speaking children (and adults) could apply prosodic information to recognise

focus in sentences. Each participant would either hear a subject- or object-accented sentence before being asked to verify the information of the pictures in terms of the sentences they heard. Contrary to Chen's (1998) findings, our results showed that Mandarin-speaking children were not sensitive to the prosodic information. Instead, they followed the word-order information – i.e., the default focus position – to verify the sentence (Xu, 2004). This was evidenced by there being more object than subject corrections in both conditions. These results provide crucial evidence that children can use specific language cues from early stages in their ambient language (in this case, Mandarin).

Both Mandarin and German speakers (adults and children) were included in the second study to confirm the assumption that children could display an adult-like performance in understanding sentence focus by identifying language specific cues in their mother tongue from an early age. In this study, we used the same paradigm (the sentence-picture verification task) as in the first study and combined this with the eye-tracking method. The visual world paradigm (VWP) has previously been used with the eye-tracking methodology to monitor how parsers comprehend ambiguous structures, and to where they direct their attention towards (Sekerina et al., 2004; Brandt-Kobele and Höhle, 2010; Höhle et al., Fritzsche and Müller, 2016). Accordingly, this method has the potential to reveal how people discover alternatives in a discourse – an ability highly suited to our research purpose. In addition to the prosodic information, we further included syntactic information (i.e., the cleft structure sentences) to the second study in order to enhance our knowledge on how Mandarin and German speakers devote prosodic and syntactic information for identifying sentence focus. First, regarding the prosodically-marked focus results, there were more object corrections in the subject-accented condition. This result is consistent with our previous findings that Mandarin speakers

apply language specific cues (i.e., word-order information) to identify the sentence focus. Second, regarding syntactically-marked focus, while Mandarin-speaking children showed an adult-like performance in the object-pseudocleft condition, they performed worse than Mandarin-speaking adults in the subject-cleft condition. However, their eye-looking patterns mirrored those of adults in the subject-cleft condition. That is, Mandarin-speaking children were aware of the syntactically-marked focus but, due to their limited processing capacities, they could not apply the information to execute a final judgement. Third, the results for the German speakers were not as clear as for their Mandarin-speaking counterparts. On the one hand, German-speaking adults performed over the chance level in the prosodically-marked focus condition, but only at the chance level in the syntactically-marked focus condition. On the other hand, German-speaking children performed over the chance level in the subject-accented, but not in the object-accented condition. Although there were two cues for marking focus in the object-accented condition (i.e., the default focus position and the prosodic information), German-speaking children may have been confused by the extra information, thereby preventing them from displaying an adult-like performance in the object-accented condition. Moreover, while they showed an adult-like performance in the subject-cleft condition, they only corrected the sentence object 26% of the time in the object-cleft condition. We assumed that German-speaking children applied a default strategy for identifying focus and further verifying the targets in the object-cleft condition. That is, they concerned the default word-order information in which focus embedded in the left periphery of sentences while correcting them. Indeed, instead of correcting the sentence object, they tended to correct the sentence subject. We thus argue that there are more possibilities of marking focus in German, such as by applying prosodic information, altering the word order, and adding

focus particles. Therefore, it would take German-speaking children longer to learn exactly how to identify focus in sentences.

We used the third study to investigate if five-year-old Mandarin-speaking children could understand the pre-subject only sentence, and whether prosodic information would help them to understand these kinds of sentences more accurately. Cross-linguistically, prior research has indicated that only children of school age could manage the pre-subject only sentences, while three-year-olds could already interpret the pre-object only sentences correctly (English: Crain et al., 1994; Person et al., 2003; German: Müller et al., 2011; Mandarin: Yang, 2002; Zhou et al., 2012; Portuguese: Costa and Szendrői, 2006). However, our results painted a different picture. We found that, with an explicit verbalised context, five-year-old Mandarin-speaking children could indeed comprehend pre-subject only sentences. Secondly, concerning, our previous research, our expectation that prosodic information would only help Mandarin-speaking adults, but not children, to more accurately understand the pre-subject only sentences was confirmed by the results. We thus found that Mandarin-speaking children (five-year-olds) could comprehend the pre-subject only sentences.

Our results seem to suggest that, from an early age, Mandarin-speaking children employ specific linguistic cues in their ambient language. That is, in Mandarin, a topic-prominent and tone language, word order information plays a more important rule than the prosodic information, and even three-year-old speakers appear able to follow the word order information. Moreover, while it seems that the German-speaking children could follow the prosodic information, they did not show a similar, adult-like performance in the object-accented condition. This could reasonably be explained by there being more possibilities of marking focus in German, meaning that mastering these linguistic tools requires more time. Another important empirical finding regarding syntactically-marked

focus in German is that it seems that the cleft construction is not a valid focus construction – a finding that accords with previous observations (e.g., Dufter, 2009). Furthermore, we found that the eye-tracking method was useful for uncovering how parsers direct their attention for recognising focus. Finally, with the help of the explicitly verbalised context, five-year-old Mandarin-speaking children could comprehend the pre-subject sentences.

Acquisition of Focus- in a Cross-Linguistic Perspective

You might not have read a book of how to speak your mother tongue while acquiring it, but you might have read a book about how to better communicate your ideas to the others when you begin your career. To communicate successfully and efficiently to the listeners, one of the important facts is how you, as a speaker, deliver your message. Before Sir Winston Churchill became the noblest speaker in the world, he studied and practiced greatly how to be an impressive speaker. In order to deliver a memorial speech to the public audience, he was competent in applying rhythms and pauses which are considered as linguistics tools. To transfer one's own idea effectively, speakers would make use of several different linguistics tools, such as additional prosodic information, the word order, definite or indefinite articles, etc. Meanwhile, the hearers would have to identify the cues to comprehend the intention of the speakers. The same holds for children: while acquiring their languages, they would need to learn how to take advantage of these linguistics cues in their own language production as well as in understanding the intention of another speaker. The study of information structure (IS) refers to the study of how the speakers package information and how they deliver information to the hearers by manipulating syntactic, prosodic, or morphological means (Chafe, 1976). In this thesis, we specifically explore how children acquire these different linguistics cues as well as whether these cues are applied differently across speakers of different languages. Last but not least, the developmental trajectory of acquiring these cues is investigated. To reach the goal, we

apply experimental methods to look into how children understand these different linguistics cues and whether there are differences between speakers of different languages.

1 Introduction

1.1 Background

How children acquire languages has been one of the most important issues in research of early human cognitive development. Understanding how language is preserved and represented by human beings will help to answer the question. More, language acquisition involves the acquisition of a complex system involving subsystems such as phonology, morphology, syntax etc. Another aspect regarding language acquisition is the acquisition of information structure and how humans encode the information structure. The study of information structure (IS) examines how the speakers organize their thoughts and deliver information to the hearers while the speakers accommodate their utterances to the temporary state of the hearer's mind (cf. Chafe, 1976- information packaging). In addition, how the information is delivered is related to the speakers' intensions and the hearers' interpretation. Therefore, to effectively process information structure, not only core linguistics competence (e.g., phonology, morphology, syntax, semantic) but also some extra-linguistics knowledge (e.g., pragmatics) is necessary. Consider the following examples (1)-(2):

1. John only gave the chocolate to SUE_F[1] (but not Mary or Rose).

2. John only gave the $CHOCOLATE_F$ to Sue (but not book or car).

---

[1] 'F', as well as capital letters, marks the focus of the sentence.

For both examples, the truth condition is that John gave the chocolate to Sue, and the difference between Example (1) and Example (2) is that in Example (1), the focus of the sentence is on the recipient, Sue (no one else but Sue), whereas in Example (2), the focus of the sentence is on the theme, the chocolate (nothing else but chocolate). English speakers can apply the accentuation on the element that they intentionally emphasize as the focus of Example (1) or (2). As a matter of fact, it might take longer for the younger learners – children - than adult speakers to effectively manage the information; more, how they acquire the information structure and manage IS becomes an essential question for linguists. Several different notions were used by linguists while examining IS across different languages., e.g., rheme and theme, topic and focus, givenness and new, etc. This thesis looks at one of the specific properties in information structure, that is focus. Before we introduce our studies, studies and properties of focus will be discussed in detail in the following.

1.1.1 Focus

Following Alternative Semantics (Rooth 1985, 1992), focus is an indication of the alternatives related to the current discourse. Consider Example (3) and (4):

3. Q: Who wrote the article *A theory of focus interpretation*?

   A: Rooth_F wrote the article. (not Jackendoff, Krifka etc.)

4. Q: Which book did Rooth write about focus?

   A: He wrote *A theory of focus interpretation*_F. (not *Focus and Presupposition*, *Basic Notions of Information Structure* etc.)

In Example (3) and (4), constituents considered as the felicitous answers to the questions, i.e., *Rooth* and *A theory of focus interpretation* are placed in the domain of focus. More,

these asserted answers contrast the other potential answers. For example, the answer in Example (3) indicates that no one else than Rooth wrote the article while the answer in Example (4) implies that Rooth wrote the article *A theory of focus interpretation* but not the other articles, e.g., *Focus and Presupposition*, *Basic Notions of Information* etc. Cross-linguistically, different tools, e.g., prosodic information, syntactic information, and lexical information, are applied for marking focus (Gussenhoven, 1984; Féry, 1993; Ladd, 1996; Vallduví and Engdahl, 1996). For languages with strict word order (e.g., English) prosodic information is considered as the primary tool for indicating focus. However, syntactic information such as the cleft construction is used for the same purpose as well. As a matter of fact, for English speakers, they could either add an accent on "Rooth" in Example 3 or they could apply the cleft construction- *it is Rooth who wrote the article*- to indicate no one else but Rooth wrote the article. In contrary, in languages with relatively flexible word order (e.g., German, Italian, Spanish) focused constituents can be moved to a privileged sentential position where focus is assigned but again accent is also available for focus indication in these languages (Italian: Frascarelli, 2000; Spanish: Zubizarreta, 1998; Feldhausen and Vanrell, 2014; German: Repp and Drenhaus, 2015).

Apart from these intonation languages, in the current study, a tone language, and a Topic-prominence language -Mandarin- will be investigated. Although Mandarin has been considered as a language with either SOV or SVO word order, this language also shows to some extent a flexible word order such that the sentence object is also allowed to move to the sentence initial position, i.e., the Topic position (for detailed discussion see Li and Thompson, 1989). In addition, as a tone language, prosodic information plays a role not only on the sentence level for indicating focus but also on the lexical level for differentiating lexical meaning in Mandarin. Due to this multifunctional operation of prosodic information in Mandarin, researchers suggested that word order information

should be more relevant than prosodic information in marking and identifying focus in a sentence (Feng, 2003; Shyu, 2012; Xu, 2004). For example, Xu (2004) proposed that once the focused element is realized in the default focus position of Mandarin, i.e., the sentence final position, prosodic focus marking is not mandatory anymore. So far, different means of representing focus in different languages have been discussed. Next, we will look into some empirical evidence of how these different means, i.e., prosodic information, syntactic information and the focus particle-only, are acquired by children in different languages.

1.1.1.1 Acquisition of Prosodically-Marked Focus

Looking at children's production, it has been indicated across different languages that children can highlight a focused element using accentuation from early on (English: MacWhinney and Bates, 1978; Wieman, 1976; Hornby and Hass, 1970; German: Weber et al., 2018; Dutch: Chen, 2011; Mandarin: Yang and Chen, 2014). However, comprehension studies have found that children up to school age still had difficulties in representing prosodically-marked focus in their sentence understanding (see Cutler and Swinney, 1987). For example, a study by Wells and her colleagues (2004) showed that English-speaking children could mark focus by using prosody information from an early age onwards in their production (in their case, five-year-olds), while children up to ten years of age still could not comprehend prosodically-marked focus correctly. This finding indicates that the acquisition of prosodic means of focus marking may be earlier in language production than in language comprehension.

Much research has been conducted on intonation languages, but there is little research concerning acquisition of prosodically-marked focus on tone languages; hence,

this study will focus on Mandarin as a tone and Topic-prominent language (Chao, 1968; Li and Thompson, 1976). Using a sentence verification tasks, Chen (1998) found that prosodic information is more prominent than word order information for sentence comprehension of Mandarin-speaking children (5-, 7-, 9-, 11-, and 13-year-old) while adults relied more strongly on word order than on prosodic information. In addition, some production studies (Chen and Braun, 2006; Chen and Gussenhoven, 2008; Ouyang and Kaiser, 2014; Xu, 1999) indicate that Mandarin-speaking children can apply prosodic information to mark focus. For example, Yang and Chen (2014) revealed that even four-year-old Mandarin-speaking children could produce focused element by making use of prosodic information (pitch and duration). In contrast to the findings with intonation language speakers, these results suggest that Mandarin-speaking children at age five, the tone language learners, could understand prosodic information as marking focus as well as producing focus by using prosodic information (Comprehension study: Chen, 1998; Production study: Yang and Chen, 2014). But again, the issue of comprehending prosodically-marked focus with Mandarin speakers needs to be further explored concerning the small amount of the related research: Chen's study as far as we know was the only one examining whether Mandarin-speaking children were sensitive to the prosodically-marked focus in their sentence comprehension. More, there is a discrepancy between children' results and adults' results regarding understanding prosodically-marked focus in Mandarin.

A study by Szendrői, Bernard, Berger, Gervain and Höhle (2017) followed the hypothesis that children's performance in tasks that require the exploitation of accentuation as focus marker would be strongly affected by the experimental design used in the respective study (see also Cutler and Swinney, 1987). Hence, in their study, they developed a new sentence-verification task to examine whether three- to six-year-old

German, English, and French-speaking children could understand prosodically-marked focus. Their results showed that all child participants reached adult-like performance across the different languages. Therefore, in order to shed more light on understanding how tone language speakers, i.e., Mandarin speakers, represent prosodically-marked focus, the same experimental task based on Szendrői and her colleagues' design will be implemented for the current study (see Chapter 2). On the other hand, Cutler and Swinney (1987), assumed that children could not reach adult-like performance in understanding prosodic information if the conducted task requires additional cognitive demands. Therefore, they suggested that instead of off-line measurements, on-line tasks would be more suitable for investigating how speakers (and specifically children) exploit prosodic information. Following Cutler and Swinney, a fine-grain online method, i.e., the eye-tracking method, will be incorporated into the experiments of this thesis together with the sentence-verification task designed by Szendrői and her colleagues (2017). It will be investigated how Mandarin and German speakers comprehend sentences with prosodically-marked as well as with syntactically-marked focus (see Chapter 3).

1.1.1.2 Acquisition of Syntactically-Marked Focus

Much less research has investigated the acquisition of syntactically-marked focus (Höhle et al., 2016). In addition to the prosodic cues, focus can be marked by syntactic cues, e.g., using non-canonical word order or using the cleft construction. A speaker prefers to use a cleft construction (Examples 5a and 5b), rather than a simple transitive sentence (Example 6) in specific discourse contexts; for example, when 'Rooth' is new or contrastive information.

5. (a). It was Rooth who wrote the article.   *It-cleft*

(b). Who wrote the article was Rooth.   *Pseudocleft*

6. Rooth wrote the article.

Using a sentence-picture verification task, Hupet and Tilmant (1989) investigated whether four- to ten-year-old French-speaking children can produce the cleft construction spontaneously when the functional necessity was granted by the context. Their results showed that while children applied the cleft construction to correct the agent, accentuation was in use for correcting the patient of a sentence. This result was different from an earlier report by Hornby and Hass (1970) that English-speaking preschoolers only used accentuation but not the it-cleft construction to convey new/contrastive information regardless of the thematic roles, agent or patient. Concerning children learning Mandarin, Chen (1998) found that only children from eleven years onwards could better comprehend the cleft construction than younger peers by using the sentence-verification task. Overall, regarding the little amount of the research on this topic, there is a need to further examine the area and we will investigate how Mandarin-speaking and German-speaking children understand the cleft construction in Chapter 3.

1.1.1.3 Acquisition of focus particle – Only

Apart from prosodically-marked focus and syntactically-marked focus, another linguistics means, the focus particle- only- is also used for highlighting the focused element. Cross-linguistic research has been shown that children up to school age had difficulties in understanding the sentences with the focus particle- only (English: Crain et al., 1994; Person et al., 2003; German: Müller et al., 2011; Mandarin: Yang, 2002; Zhou et al., 2012; Portuguese: Costa and Szendrői, 2006).

Consider the Example (7):

7. Rooth only wrote the article.

    a. Rooth only wrote the [article]F. (Rooth only wrote the article, and nothing else was present in the discourse model.)

    b. Rooth only [wrote the article] F. (Rooth was only writing the article and no other activity was present in the discourse model.)

Following Rooth (1992), focus indicates the alternative representing in the current discourse. One feasible representation of Example (7) is that Rooth only wrote the article and nothing else than the article as shown in (7a), or it could be interpreted as in (7b) that Rooth was only writing the article and he wasn't doing any other activities. Again, as mentioned above, one way to disambiguate between these two interpretations is to place the accentuation upon the focused element, e.g., either on "the article" or on "wrote the article". In this project, we further look at whether the prosodic information will help tone language learners- Mandarin-speaking children- to better comprehend the sentences with the focus particle- only in Chapter 4.

1.1.1.4 Current state of research in acquisition of prosodically-marked focus and syntactically-marked focus

Overall, there is much more research on children's acquisition of prosodically-marked focus in intonation languages than in tone languages. The same holds for the research on syntactically-marked focus. Therefore, to further compare how intonation language speakers and tone language speakers identify prosodic information and syntactic information for marking focus, the present study investigates how Mandarin speakers (adults and children) use prosodically-marked focus as well as syntactically-marked focus

by using experimental methods. Further, the same experimental design is applied with speakers of German as an intonation language (adults and children). Finally, we explore whether Mandarin-speaking children could benefit from the prosodic information to better understand sentences with the focus particle, *only*.

The following research questions are addressed in this thesis:

1. How do children and adults exploit prosodic and syntactic information to understand focus across different languages, i.e., in Mandarin and in German?

2. Is there a difference between children's and adults' performance concerning processing focus? If there is, how different are they?

3. Would an online method, i.e., the eye-tracking methodology, help us to discover how Mandarin-speaking children and German-speaking children process the prosodic cues as well as the syntactic cues of marking focus?

1.2 Outline of the study

The organization of the thesis is as follows. In Chapter 1, the central issue of language acquisition concerning information structure, specifically literature on acquisition of different types of focus will be discussed. In Chapter 2, a study will be described that approaches the vital question of whether Mandarin speakers can exploit prosodic cues of focus marking for sentence comprehension. A sentence-verification task with three to five-year-old and adult Mandarin speakers was used in this study. Further, in Chapter 3, in addition to the prosodically-marked focus, a study concerning how German and

Mandarin speakers (adults and children) apply syntactic cues as well as prosodic cues to identify focus in sentences will be reported. Chapter 4 will provide evidence of how Mandarin speakers (adults and children) understand sentences with the focus particle 'only' which need the consideration of prosodic information. Last but not least, Chapter 5 will present a summarizing discussion of the findings of the studies that have been presented in the previous chapters and provide some considerations on future outlooks for this research field.

2 Understanding prosodic focus marking in Mandarin Chinese-Data from children and adults²

2.1 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]F³ 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

---

³ F means the focus of the sentence.

privileged sentential positions. Nevertheless, this typically goes hand in hand with prosodic highlighting of the focused constituent (Italian: Frascarelli, 2000; Spanish: Zubizarreta, 1998; Feldhausen and Vanrell, 2014; German: Repp and Drenhaus, 2015).

Across languages, a focused constituent is often associated with prosodic salience, where salience is achieved by using phonetic properties including pitch, 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 multiple 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 precisely reached 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 and Braun, 2006; Chen and Gussenhoven, 2008; Ouyang and 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 favoured 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 behavioural 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.

2.2 Method

2.2.1 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[4]. 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 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.

---

[4]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.

2.2.2 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
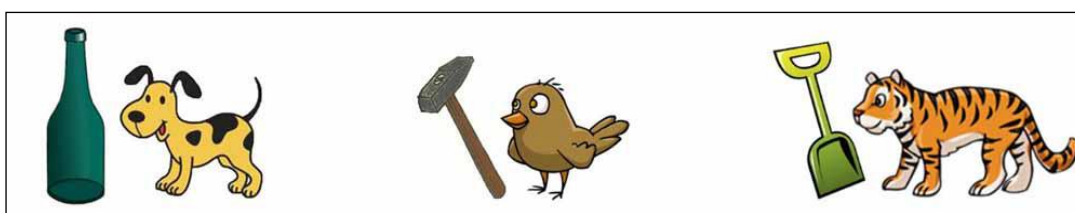
Table 1 Examples of the test sentences used as target and control trials5

Subject-accented sentence	Object-accented sentence
Example1: Target Trial	**Example2: Target Trial**
XIAONIAOF you shueping, shi ma?	xiaoniao you SHUEPINGF, shi ma?
BIRDYF has bottle, Aux Q?	birdy has BOTTLEF, Aux Q?
'The BIRDYF has the bottle, is that right?'	'The birdy has the BOTTLEF, 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.
Example 3: Control Trial	**Example 4: Control Trial**
XIAONIAOF you chuizi, shi ma?	xiaoniao you CHUIZIF, hi ma?
BIRDYF has hammer, Aux Q?	birdy has HAMMERF, Aux Q?
'The BIRDYF has the hammer, is that right?'	'The birdy has the HAMMERF, 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

---

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

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).

2.2.3 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.

2.2.4 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.

2.3   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 condition6.

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 (Group3&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 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.

6Response~group*condition+(1|id)+(1|item)

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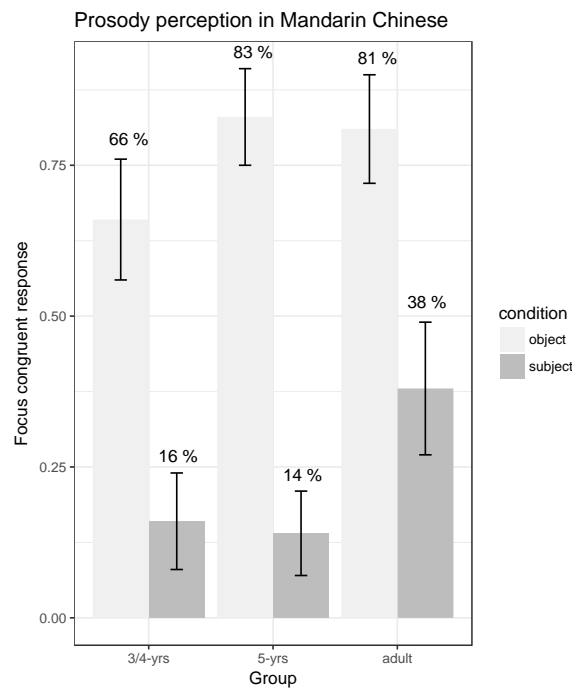
2.4   Discussion

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 and Höhle, 2012; Papafragou and 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.

3 Comprehension of focus structures in a cross-linguistic perspective—Evidence from five-year-old Mandarin and German-learning children

In the previous chapter, we have investigated whether Mandarin speakers (adults and children) were sensitive to prosodically-marked focus. The results demonstrated that even younger children (i.e., three-year-olds) did not have this sensitivity. Instead, for Mandarin speakers, word order information played a more important role in understanding focus in sentences. In addition to prosodically-marked focus – the subject- and the object-accented condition – this chapter will explore syntactically-marked focus by using cleft constructions. Further to Mandarin (a tone language), we will also here examine German (an intonation language). This cross-linguistic study seeks to contribute to a more comprehensive understanding of whether different languages speakers would distinctly comprehend disparate focus marking cues, and whether five-year-olds apply syntactic and prosodic cues of focus marking in a similar way as adult speakers in their ambient languages. As research regarding children's acquisition of prosodically-marked focus has already been reviewed in Chapter 2, this chapter is organised as follows: Section 3.1 will present only relevant previous research on the cleft construction, followed by a brief research review on eye-tracking experiments as this method has been used in the chapter's experimental study. We will present our hypotheses before introducing the current experimental design and the results in Mandarin (Section 3.2) and in German (Section 3.3). That is, German speakers would perform better in the prosodically-marked focus condition than in the syntactically-marked focus condition, whereas Mandarin speakers would be sensitive to the syntactic, rather than prosodic, cues. Finally, we will

summarise and compare the results in Mandarin and German and provide a discussion on future research avenues.

3.1 Introduction

The sentence structure of the English cleft construction embodies both a copula and relative clause. It has a discourse function, which serves to draw attention from the speakers to the listeners (Jesperson, 1949; Jackendoff, 1972), such as:

(1) It was JOHN[7] who ate the chocolate.

In the above example of an English cleft sentence, *John*, the so-called clefted constituent, is contained in the copula clause (for a more recent syntactic analysis, see Reeve, 20110). For this example, the discourse function of the clefted constituent *John* is the focus; however, the discourse function of the clefted constituent (i.e., focus or topic) has been disputed cross-linguistically (Prince, 1978; Declerk, 1984; Lambrecht, 1996, 2001; Hedberg, 1990, 2000; Huber, 2004). Additionally, it has been argued that, across languages, the cleft construction is associated with the exhaustiveness interpretation. Therefore, example (1) should be interpreted such that only John – and no one else – ate the chocolate (Kiss, 1998; Hedberg, 2000)[8]. In Mandarin, the cleft construction is an umbrella term for at least four constructions, i.e., the sentence initial bare *shi*, the sentence medial bare *shi*, the sentence medial *shi...de*, and the sentence initial *shi...de* constructions (for detailed discussions of these, see Tang, 1983; Cheng, 2008; Paul and Whitman, 2008; Hole, 2011). In order to be comparable to Chen's (1998) study, we included the initial

---

[7] The capital letters indicate the sentence's focus; a technique which will apply to the rest of the examples in this chapter.

[8] The exhaustiveness condition is recognised as one of the most distinguished features of the cleft construction. However, Dufter (2009) found some counter examples in his English corpus that certain clefts could accord with the additive particle 'too', thereby leading to the cancellation of the exhaustiveness interpretation.

bare *shi* construction within our research. In Mandarin, the focus in a bare *shi* cleft construction is placed immediately after the copula—*shi*—, as in:

(2) *Shi YUEHAN che le   qiaokeli.*

    Be JOHN   ate PFV chocolate.

    '*It was JOHN who ate the chocolate*'. (Subject-cleft construction)

In (2), '*John*' is in focus and immediately follows '*shi*', meaning that no one but John ate the chocolate. The function of *shi* is regarded as the focus marker comparable to *it is* in the English it-cleft construction (Teang, 1979). The corresponding sentence without *shi* (i.e., *Yuehan che le qiaokeli)* would be interpreted as having object focus by default. Moreover, focus in the bare *shi* cleft construction shifts from the default focus position – the final position of the most embedded clause (Xu, 2004) – to the front of the sentence.

In contrast to cleft constructions in such languages as English or German, there are no object-clefts in Mandarin (see Example 3(a), and for detailed discussions of the object-cleft in Mandarin, see Paul and Whitman, 2008; Hole, 20112). One way of focusing the sentence object with a syntactically-marked structure in Mandarin is through the use of the pseudocleft construction (Tang, 1980). This construction could be exploited for marking subject focus (subject-pseudocleft) or object focus (object-pseudocleft). Following Tang (1983), the current study will only consider the object-pseudocleft, such as in:

(3) a. *Shi QIAOKELO   Yeuhan che de.

      Be CHOCOLATE John   eat DE.

      'It is chocolate that John ate'. (Object-cleft construction)

    b. *Yuehan che de shi QIAOKELI.*

      John   eat DE be CHOCOLATE.

      '*What John ate was the chocolate*'. (Object-pseudocleft construction)

      c. *Che le   quiaokeli de shi YUEHAN*.

      Eat PFV chocolate DE be John.

      'Who ate the chocolate was John'. (Subject-pseudocleft construction)

While (3a) is ungrammatical in Mandarin, (3b) is correct where the object '*chocolate*' is syntactically-marked as focus by the pseudocleft construction. In example (3b), '*chocolate*', the object, follows '*shi*' immediately and appears at the final position of the sentence, resulting in a doubly-marked focus by two means. First, the focus constituent is placed in the default focus position and, second, it follows the syntactic cue *shi* (Cheng, 2008)[9]. Moreover, the pseudocleft construction can be used for marking sentence subject in focus (see Example 3c). Accordingly, it could be argued that it would be easier for Mandarin-speaking children to understand focus in the object-pseudocleft construction than in the subject-cleft construction.

Chen (1998) addressed the issue of how Mandarin speakers of all ages use prosodic and syntactic cues to interpret given and new information in four types of constructions, i.e., the canonical active, the passive, the subject-cleft, and the object-pseudocleft constructions. We have previously reviewed Chen's study (see Chapter 2) for the canonical sentences with prosodic information in which Mandarin-speaking children performed more highly in the object-accented condition than in the subject-accented condition, while the reverse was true for Mandarin-speaking adults. Her results regarding the subject-cleft and the object-pseudocleft constructions will be discussed in greater detail here. The Mandarin participants performed better with the object-pseudocleft construction than the subject-cleft construction in all age groups (five-, seven-, nine-, eleven-, thirteen-year-old, and the adults). Moreover, the participants' performance was

---

[9] Unlike the function of *shi*, the function of *de* in the Mandarin cleft and pseudocleft construction has been the subject of some debate (i.e., the sentence final particle, the past tense morpheme, an assertion operator, and a generalized λ-operator) within the literature (Cheng, 2008; Simpson and Wu, 2002). As this is exceeds this paper's scope, we shall not delve into the details of the discussion here.

increased when the accent was added to a constituent imbedded in the focus domain (Example (4a) and (4b)). However, this was not the case when the accent was added to a constituent that was not imbedded in the focus domain for the object-pseudocleft sentence and the subject-cleft sentence (Example (5a) and (5b))[10].

   (4) a. *Nanhai fangzou de shi HUDIE*$_{F(ACCENTED)}$ *ma*?

      Boy   release DE Be BUTTERFLY$_{F(ACCENTED)}$ Q

      'What the boy released was the butterfly, right?'

      (Object-pseudocleft with the accent on object)

      b. *Shi NANHAI*$_{F(ACCENTED)}$ *fangzou Hudie*   *ma*?

      Be   BOY$_{F(ACCENTED)}$ release Butterfly Q

      'It was the boy that released the butterfly, right'

      (Subject-cleft with the accent on subject)

   (5) a. *NANHAI*$_{accented}$ *fangzou de shi HUDIE*$_F$ *ma?*

      BOY$_{accented}$ release DE Be BUTTERFLY$_F$ Q

      'What the boy released was the butterfly, right?'

      (Object-pseudocleft with the accent on subject)

      b. *Shi NANHAI*$_F$ *fangzou HUDIE*$_{accented}$ *ma?*

      Be   BOY$_F$ release BUTTERFLY$_{accented}$ Q

      'It was the boy that released the butterfly, right?'

      (Subject-cleft with the accent on object)

That is, when the prosodic information was in competition with the syntactic cues, the Mandarin participants preferred the latter.

     Researchers have previously examined exhaustivity properties of the cleft construction in the acquisition of Western Germanic or Romance languages (for German

---

[10] Focus of the subject- and object-cleft, e.g., 'the boy and the butterfly' was, on the one hand, syntactically marked by the copula *shi*. On the other hand, they were prosodically marked by the accent.

and English, see Heizmann, 2007; for French, see Tieu and Križ, 2017). However, little academic attention has been directed towards the acquisition of comprehending the cleft construction. The very first study in relation to this acquisition was conducted by Lempert and Kinsborne (1980). They examined four structures – the active, the passive, the subject cleft, and the object cleft construction – in English-speaking children (from two-and-a-half-year-olds to six-year-olds) by using an act-out task. For the purpose of our study, we shall only report the results of the subject clefts (e.g. *It's the cow that bumps the horse*) and the object clefts (e.g. *It's the truck that the wagon bumps*). Across all age groups, the participants found the object clefts more difficult to understand than the subject clefts: children above the age of four-and-a-half already showed ceiling performance for the subject clefts, yet only children aged five and above could understand the object clefts (mean percentage of accuracy: 83%). Moreover, the authors assumed that children would make use of two strategies to comprehend different structures: an agent-action approach and the surface structure information strategy. In the former, children would take the noun, which is closer to the verb, as the agent of the sentence. According to this strategy, they would correctly act out the object clefts but not the subject clefts. Regarding the latter, children would take the first noun as the agent and accordingly correctly act out the subject clefts but not the object clefts. Interestingly, that the better performance with subject, compared to object, clefts aligns with previous findings on the acquisition of relative clauses: namely that, for children, the subject relative clause is easier to cross-linguistically acquire than the object relative clause (de Villiers et al., 19789; Corrêa 1995; Friedmann and Novogrodsky 2004; Diessel and Tomasello, 2005; Hu et al., 2016).

As mentioned above, there is only little previous research on the German cleft construction. Dufter's (2009) corpus study examined 459 translation counterparts of

English it-clefts in German, French, Italian, and Spanish[11]. He found that, in total, only 39 English it-cleft sentences were translated in German by using the cleft construction, while the remaining 420 were translated by using other sentence constructions. He suggested that, unlike other European languages (Portuguese, French, etc.), there were other feasible means for marking focus in German instead of using the cleft construction, such as varying word order or adding focus particles. To the best of our knowledge, no empirical study exists as to how, or whether, German-speaking adults and children understand subject and object clefts as syntactical vehicles of focus marking. Accordingly, we will address this issue by using an eye-tracking experiment with the visual world paradigm (VWP). This technique will be introduced in the following section.

It is widely agreed on that rapid real-time linguistics processing, such as syntactic parsing, spoken word recognition, etc., can be observed by recording eye movements to the referents of the words mentioned while listening to a linguistic stimulus – known as the VWP (Tanenhaus et al., 1995). Moreover, the VWP has been used to monitor how the parser computes a temporally ambiguous structure (for adults, see Tanenhaus et al., 1995; Snedeker and Trueswell, 2003; for children, see Trueswell et al., 1999; Snedeker and Trueswell, 2004). Additionally, it is widely accepted that, for adults and (young) children alike, where the eyes fixate is where people direct their attention towards (Sekerina et al., 2004; Brandt-Kobele and Höhle, 2010; Höhle et al., 2016). A sentence-verification study with the VWP by Höhle and colleagues (2016) investigated whether four-year-old German-speaking children could correctly represent pre-subject only sentences. They found that, while the children could not do so correctly (indeed, they performed below the chance level in the sentence verification task), their eye looking patterns revealed an

---

[11] Dufter (2009) used the EUROPARL corpus, a parallel corpus of European Parliament Proceedings from 1996 to 2006. It is a multilingual database containing 10 official EU languages: English, Dutch, German, Danish, Swedish, French, Italian, Spanish, Portuguese, Finnish, and Greek. Each of these languages have their counterpart translations in the other nine.

adult-like performance. They found that the children paid significantly more attention to the subject alternative in the pre-subject sentence condition than in the other two conditions (the pre-object only sentence and without focus particle sentence conditions). Similarly, the purpose of the current study is to examine whether Mandarin and German speakers (adults and children) take prosodic and syntactic cues as a means to recognise focus. In fact, the eye-tracking method matches our research purpose: eye fixations would help us apprehend where Mandarin and German speakers draw their attention to while listening to the test sentences with the VWP.

The current study

The present study seeks to understand more comprehensively how prosodic or syntactic cues are used as means to mark focus cross-linguistically by using a sentence-picture verification task with the VWP. In the previous chapter, the results showed that both adult Mandarin speakers and Mandarin-learning children were not sensitive to prosodics cues, but that the word-order cue played a more important role. In contrast, Szendrői and colleagues (2017) found that German adult speakers, as well as German-learning children, apply and interpret prosodic cues as focus markers. Moreover, both studies argue that even very young children display the same behaviour as adults. That is, children manage to comprehend the language-specific ways of marking focus from an early age onwards. Therefore, we hypothesise that Mandarin adults and children would perform better in the syntactically-marked focus condition than in the prosodically-marked focus condition (i.e., the subject-cleft and the object-pseudocleft condition vs the subject-accented condition). Conversely, we predict that German adults and children would perform more highly in the prosodically-marked focus condition than in the syntactically-marked focus condition

(i.e., the subject-accented and the object-accented condition vs the subject-cleft and the object-cleft condition.) The reason for examining different experimental conditions in Mandarin and German was due to the non-existence of the object-cleft construction in Mandarin. Therefore, instead of the object-cleft construction, the object-pseudocleft construction was tested for Mandarin speakers. We included not only an explicit method – a sentence-picture verification task – but also an implicit method – the eye-tracking task. Applying the eye-tracking method could help us gain deeper insights into how different speakers parse the focus structures.

3.2 Study with Mandarin speakers

3.2.1. Method

3.2.1.1. Participants

52 five-year-old children (mean age: 5.07, age range: 5.01–5;12) participated in the study. All were reported as typically developing children by their parents. They were randomly divided into three groups, each of which group was tested with one experimental condition: 18 in the subject-cleft condition, 17 in the object-pseudocleft condition, and 17 in the subject-accented condition. Additionally, 56 (under)graduate students at National Chengchi University served as the control group[12], and were also randomly divided into three subgroups: 17 in the cleft-subject condition, 19 in the pseudocleft-object condition, and 20 in the subject-accented condition. All were native speakers of Mandarin Chinese

---

[12] All students began learning English as a foreign language at the age of 12 and had infrequent contact with English at the time of testing.

living in Taipei, Taiwan, and had no known visual or audio impairments at the time of testing.

3.2.1.2. Materials and design

Sentence-picture verification task with the VWP. The experiment consisted of 23 trials. These comprised an introduction trial, two practice trials, four filler trials, eight test trials, and eight control trials. The trials were pseudo randomised so that no two same trial types would follow consecutively, and two different order lists were created.

To investigate whether Mandarin-speaking children could apply prosodic or syntactic cues to understand focus, we adapted the sentence-picture verification task from Chen et al.'s (2018) study, in which participants were asked to correct statements about the contents of a visual display they were currently viewing. We examined the subject-accented condition (see Chapter 2) and two syntactically-marked focus conditions (i.e., the subject-cleft condition and the object-pseudocleft condition). In the subject-accented condition (Examples 1&2 in Table 1.) and in the subject-cleft condition (Examples 3&4 in Table 1.), the subject was the focus of the target sentence, whereas focus was placed on the object of the target sentence in the object-pseudocleft condition (Examples 5&6 in Table 1.).

Table 1. Examples of the target and control trials for the three conditions

Target Trial	Control Trial
Example 1: Subject-accented condition	**Example 2: Subject-accented condition**
XIAONIAO$_F$[13] you shueping, shi ma?	XIAONIAO$_F$ you chuizi, shi ma?

---

[13] See Footnote 1.

THE BIRDY$_F$ has bottle Aux Q?	THE BIRDY$_F$ has hammer, Aux Q?
'THE BIRDY$_F$ has the bottle, right?'	'THE BIRDY$_F$ has the hammer, right?'
Focus-congruent response:	Focus-congruent response:
Bushi, xiaogou you shueping.	Shi.
'No, the dog has the bottle.'	'Yes.'
Example 3: Subject-cleft condition	**Example 4: Subject-cleft condition**
Shi XIAONIAO$_F$ you shueping, shi ma?	Shi XIAONIAO$_F$ you chuizi, shi ma?
Be BIRDY$_F$ has bottle Aux Q?	Be BIRDY$_F$ has hammer Aux Q?
'It is the BIRDY$_F$ who has the bottle, is that right?'	'It is the BIRDY$_F$ who has the hammer, is that right?'
Focus-congruent response:	Focus-congruent response:
Bushi, shi xiaogou you shueping.	Shi.
'No, it's the dog who has the bottle.'	'Yes.'
Example 5: Object-pseudocleft condition	**Example 6: Object-pseudocleft condition**
Xiaoniao you de shi SHUEPING$_F$, shi ma?	Xiaoniao you de shi CHUIZI$_F$, shi ma?
Birdy has de be BOTTLE$_F$ Aux Q?	Birdy has de be HAMMER$_F$ Aux Q?
'What the birdy has is the BOTTLE$_F$, is that right?'	'What the birdy has is the HAMMER$_F$, is that right?'
Focus-congruent response:	Focus-congruent response:
Bushi, xiaoniao you de shi chuizi.	Shi.
'No, what the bird has is the hammer.'	'Yes.'

We used the pictures from Szendrői et al.'s (2017) study in our own. The materials were modified so that they could be applied in the VWP. In addition to three pairs of

animals and objects equally distributed across the picture (left, middle, and right), we included a girl (as the experimenter) hiding behind a board (See Figure 1.). In the subject-accented condition and the subject-cleft condition, the animal which corresponded to the constituent in focus was always placed in the middle of the three animal-object pairs. In the object-pseudocleft condition, the object which corresponded to the constituent in focus appeared four times on the left-hand side of the display and four times on the right-hand side. In the eight target test trials, the descriptions of the pictures were always false, whereas the test sentences of the control trials were true descriptions of the corresponding pictures – presupposing that the listener was sensitive to the prosodic or syntactic focus marking. The purpose of the control trial was to balance out the negative and positive responses as all test trials were intended to evoke a negative response. Moreover, the construction of the filler trials was similar to the test and control trials, but differed in that a sentence of a filler trial highlighted a shared feature among the three animals or objects. For example, a test sentence of a filler trial corresponding to Figure 1. could be 'Are all animals green?'. An expected answer would be 'No, the doggie is yellow, the birdy is brown, and the tiger is orange'. There were two false and two true statements among the filler trials. The purpose of the filler trials was to ensure that the participants would check all of the animals and objects before making any judgments. As such, only participants who were correct in the filler and control trials would be included in the data analysis of the test trials. Parents were asked to complete a parental questionnaire after the test to confirm whether the children were familiar with the animals and items used in the task. All test sentences were pre-recorded by a female Mandarin speaker. To run the experiment on the eye-tracker, all images were combined with the corresponding audio stimulus into video clips.

Figure 1. Example of an experimental image including three pairs of objects and animals, and an experimenter with a board in the background. The size of each white area was controlled (300x300 pixels).
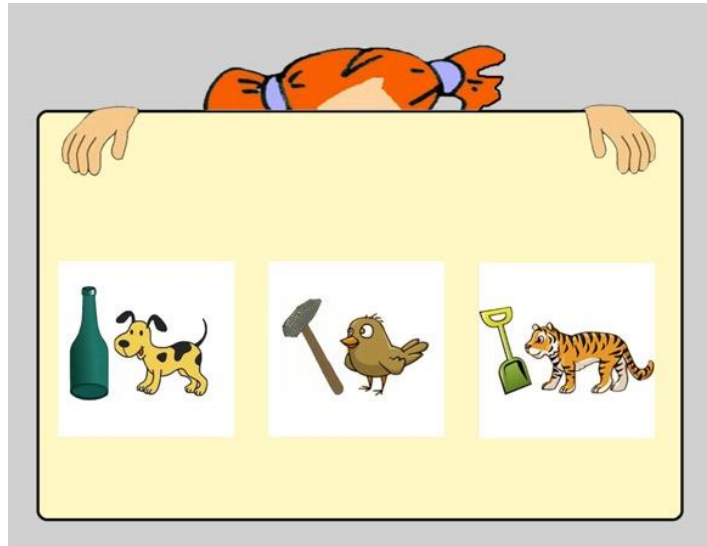
3.2.1.3. Procedures

The eye gaze data and verbal responses were collected in the experiment. To collect the data, we used the Tobii Pro TX300 – a binocular corneal reflection eye tracker, incorporating a 23-inch Monitor, which tracks the eye gaze positions approximately every 3ms with a tracking rate of 300 hertz. The verbal responses were first video-recorded and later manually annotated. Experimental materials were presented with Tobii studio (version: 3.0.4) which was applied to simultaneously record the eye gaze data. The participants were seated approximately 60–70cm in front of the eye-tracker. The experiment started with a five-point calibration programming by the Tobii Studio. The experimenter judged the quality of calibration in accordance with the criteria of the Studio software. If the eye-tracker was found to correctly track the participants' eyes, the task would start immediately. The experiment began with an introduction trial, in which

the participants were informed that a girl was trying to remember the pictures shown during the task. The girl would describe the pictures currently being viewed, and the participants were requested to verify the statements. If she was wrong, the participants were asked to correct the statement. Two practice trials followed the introduction trial. Every trial, except the introduction trial, started with the presentation of the visual stimulus (which lasted 10–11 seconds in total). The audio stimulus began one second after the appearance of the visual stimulus. Audio stimuli had a duration of approximately 2–3 seconds. The average length of the audio duration in each condition was the following: 2.789s in the subject-accented condition (SE: 0.217), 2.067s in the subject-cleft condition (SE: 0.107), and 2.203s in the object-pseudocleft condition (SE: 0.104). Participants were asked to judge each test sentence and give a yes-no response and correction with an SVO sentence as soon as possible after the end of the audio stimulus (examples for each condition can be found in Figure 2). All participants were tested in a quiet room either at the university or in the kindergarten.

Figure 2. Illustration of an experimental trial

Trial starts (Eye movements recording starts)

1000ms (silence viewing)

The target sentence begins

2000ms-3000ms

End of the target sentence

7000 ms-8000ms
(Meanwhile the participant verifies the target sentence)

End of the trial
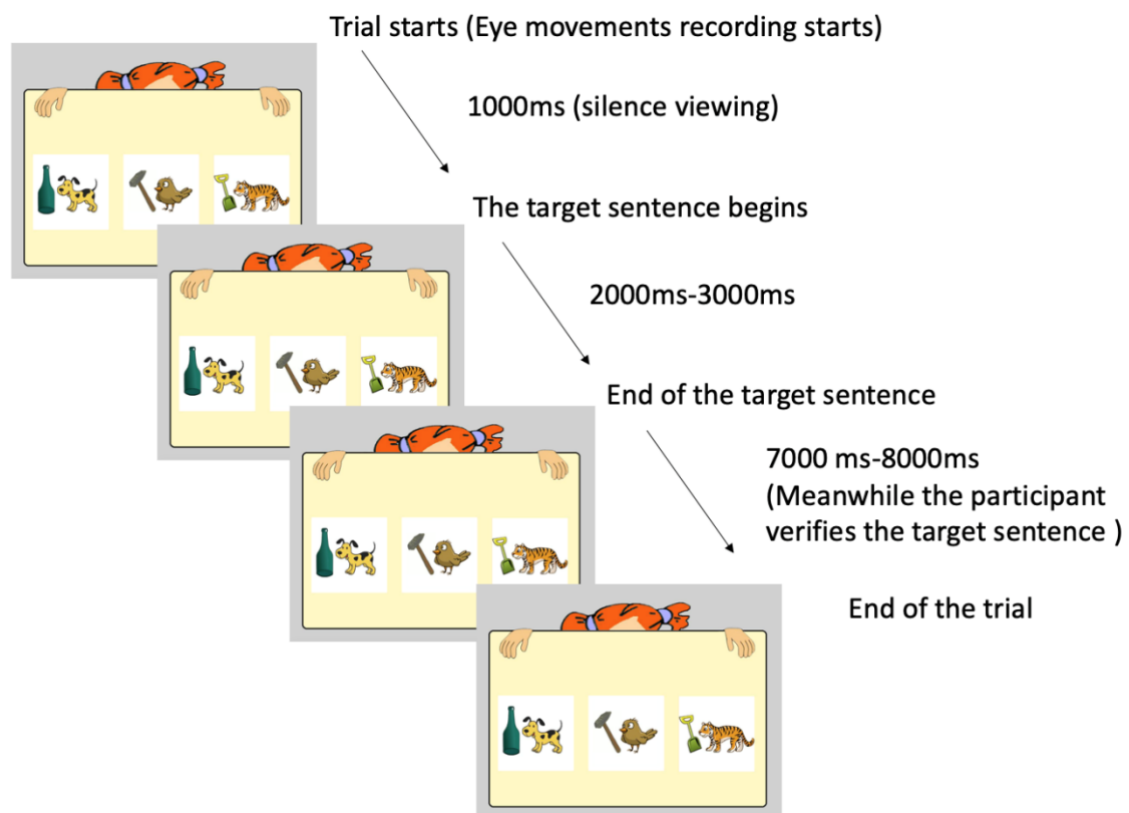
3.2.1.4. Scoring of the response data

Corrections of the subject in the subject-accented and subject-cleft conditions, and corrections of the object in the object-pseudocleft condition, were considered focus-congruent responses and scored as correct. However, corrections of the object in the subject-accented and subject-cleft conditions, and corrections of the subject in the object-pseudocleft condition, were considered to be focus-incongruent responses and scored as incorrect. Additionally, there were 14 cases from the child participants and 11 cases from the adult participants in the subject-cleft condition (5% in the children's response data and 4% in the adults' response data), and 4 cases from the adult participants in the object-pseudocleft condition (1% in the adults' response data) where both the subject and object

were corrected at the same time (e.g., 'No, the doggie has the bottle and the birdy has the hammer'). We counted these responses as invalid and excluded them from the statistical analysis. Moreover, there were cases where participants used a phrase, rather than a full sentence, as a response, such as the subject correction 'the dog' in the subject-accented and subject-cleft conditions, or the object correction 'the hammer' in the object-pseudocleft condition. However, these were still coded as valid and correct responses. On the other hand, the object correction 'the hammer' in the subject-accented and subject-cleft conditions, and the subject correction 'the bird' in the object-pseudocleft condition were coded as valid but incorrect responses.

3.2.2. Results

3.2.2.1. Verbal responses

All responses for the filler and the control trials were correct, meaning that no participants were excluded from statistical analysis. For inferential statistics, we used a linear mixed-effect model (Baayen et al., 2008) in R (R Core Team, 2016) in the R-studio environment (version 1.1.383). The aim was to analyse the impact of Sentence Conditions and Age Groups, as well as their interactions, on the number of focus-congruent responses in the test trials. Further to these fixed factors, two random effects were included in the model, in which by-subjects and by-items variance to the responses were considered (the random slope[14]. The contrast for Age Groups was established such that children was considered as a 'baseline' so as to allow a comparison between theirs and adults'

---

[14] The linear mixed-effect model used in the current study is presented here:
Response~group*condition+(1|id)+(1|item)

performance. On the other hand, the contrast for Sentence Conditions was established to have the subject-cleft as a 'baseline' in order to facilitate a comparison with the other two sentence conditions (the subject-accented and the object-pseudocleft). That is, we tested whether the number of focus-congruent responses in the subject-cleft condition in child data differed significantly from those in the adults' data. Moreover, for the child group, we tested for whether the number of focus-congruent responses in the subject-cleft condition was significantly different from the number of focus-congruent responses in the subject-accented and object-pseudocleft conditions. Finally, the interaction of Age Groups and Sentence Conditions was analysed. In so doing, we found that the differences between the subject-cleft and subject-accented conditions, and the differences between the subject-cleft and object-pseudocleft conditions for these two age groups were significantly distinctive. Finally, we employed the Wilcoxon test for the chance level tests with both participants in all conditions.

The results are shown in Figure 3. Overall, the number of focus-congruent responses of the child participants in the subject-cleft condition was significantly lower than those in the object-pseudocleft condition (Group$_{children}$: b = -7.66, SE = 2.93, Z = 5.851, p < 0.001), but not significantly different from the number of focus-congruent responses of the child participants in the subject-accented condition (Group$_{children}$: b = -1.60, SE = 2.27, Z = -0.704, p = 0.481). The children's performance in the subject-cleft condition was significantly worse than that of adults (Group$_{adults}$: b = 15.476, SE = 3.13, Z = 4.942, p < 0.001). Moreover, due to the child group's poorer performance in the subject-cleft condition, there were significant effects for the interaction of Age Group and Conditions (Age Group$_{adults\ vs\ children}$*Condition$_{subject-cleft\ vs\ subject-accented}$: b = -15.850, SE = 3.80, Z = -4.127, p < 0.001; Age Group$_{adults\ vs\ children}$*Condition$_{subject-cleft\ vs\ object-pseudocleft}$: b = -14.847, SE = 4, Z = 3.71, P < 0.001). To summarise, there were differences in the performance of children

and adults in the subject-cleft condition, but none were observed in the other two conditions (the subject-accented and object-pseudocleft conditions).

Additionally, both the child and adult group performed better than the chance level in the object-pseudocleft condition ($p = 0.003$ and $p = 0.0003$, respectively). Moreover, children performed at the chance level in the subject-cleft condition ($p = 0.202$), while the adult participants outperformed it ($p = 0.049$). Both groups performed below the chance level in the subject-accented condition (the child group: $p = 0.003$ and the adult group: $p = 0.0005$).

Figure 3. Percentage of focus-congruent responses across Sentence Conditions and Age Groups. Each error bar shows the 95% confidence intervals for the mean.
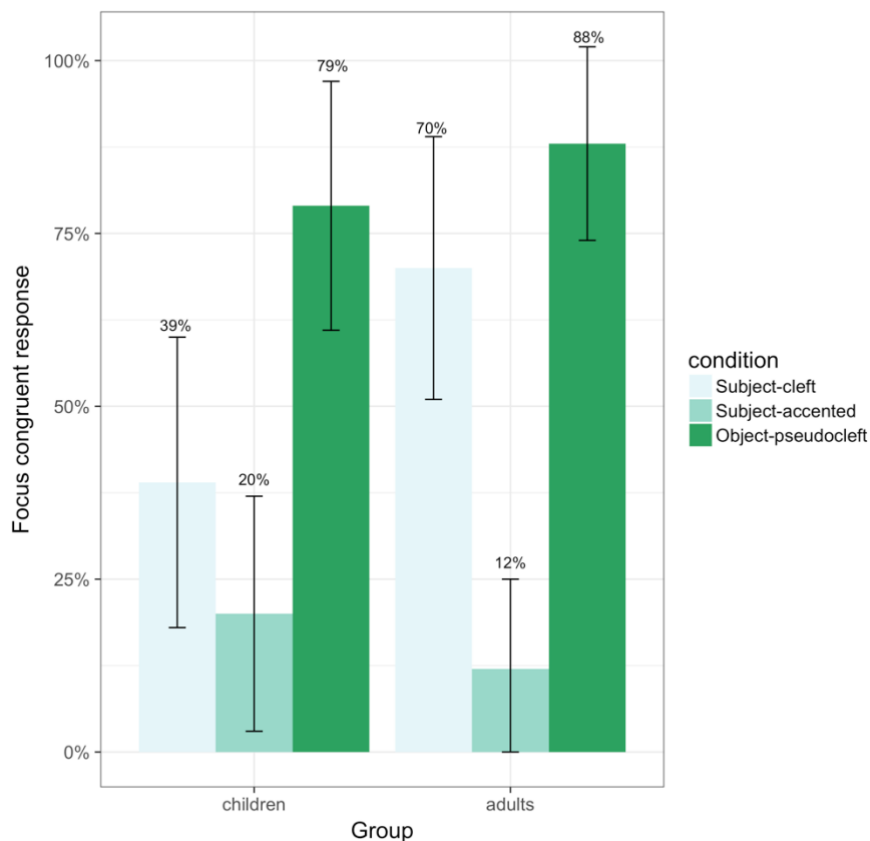
3.2.2.2. Looking pattern

3.2.2.2.1. Data treatment

The eye gaze data were extracted and processed using Tobii Studio software (version 3.0.4). For the target test trials, only the data after the end of the audio stimuli – which lasted approximately 7–8 seconds – were extracted for statistics. The reason for choosing the time window is that the participants were unaware of what to answer only until the target questions were asked. Hence, this time window was chosen to indicate how children process the prosodic or syntactic cues of focus marking after being provided with the target question. Furthermore, we sought to observe how the prosodic and syntactic information guided the participants' attention. Trials that showed a track loss of over 50% were excluded (10.5% of the adult data and 10.2% of the children's data).

3.2.2.2.2. Region of interest

The region of interest (ROI) was determined as the area on the display that showed the referent of the alternative subject NP (see Figure 4.). The alternative subject NP was positioned on the left for half of the time and on the right for the remaining half – a positional variable which was considered in the statistic model. For the subject-accented and subject-cleft conditions, glances to this ROI were considered as relating to the alternatives for the focused element. In contrast, this ROI does not contain the alternatives for the focused object in the object-pseudocleft condition. Therefore, in this condition, we expected fewer looks to this ROI, and data from this condition served as the contrast data set to that of the other two conditions.

Figure 4. Example for a visual display with the ROI is marked by a red frame (300x300 pixels). The red frame is here only included for illustration, and was not part of the display used in the experiment.
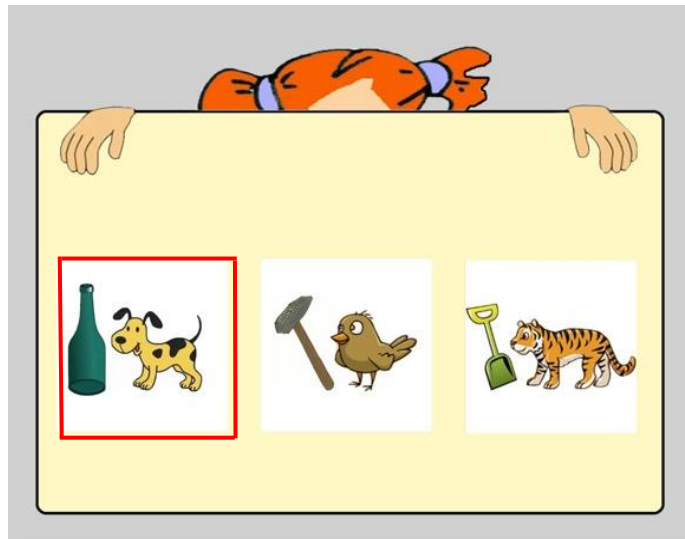
3.2.2.2.3. Data analysis

A linear mixed model (Baayen et al., 2008) in R (R Core Team, 2016) under R Studio (version: 1.1.383) was used. The aim was to evaluate the impact of Sentence Conditions and Age Groups, and their interactions, on the proportions of looks to the ROI. In addition to these fixed factors, we included three random effects, in which by-subjects, by-items, and by-position variance (the position of the ROI) to the proportions of looks to the ROI were taken into account (the random slope)[15]. Again, to be comparable to the model for the response data (see Section 3.2.2.2.), the contrast for Age Groups considered children as the 'baseline' in order to compare their performance with that of adults. For the Sentence Condition factor, the subject-cleft condition was set as the 'baseline' so as to

---

[15] Proportions of looks to the ROI~group*condition+(1|id)+(1|item)+(1|position)

form a comparison with the subject-accented and object-pseudocleft conditions. That is, we tested whether the proportions of glances to the ROI in the subject-cleft condition differed significantly between age groups. Moreover, within the child group, we tested whether the proportions of looks to the ROI in the subject-cleft condition was significantly different from those in the subject-accented and object-pseudocleft conditions. Furthermore, for the interaction, the effects of the differences between the subject-cleft and subject-accented conditions, and the effects of the differences between the subject-cleft and object-pseudocleft conditions for these two age groups were examined.

The eye-movement results are shown in Figure 5. As with the verbal response data model, the children's data in the subject-cleft condition was used as the baseline. The proportion of looks to the ROI in the subject-cleft condition of the children was significantly higher than in both the subject-accented condition (Group$_{children}$: b = -0.253, SE = 0.049, Z = -5.151, p < 0.001) and the object-pseudocleft condition (Group$_{children}$: b = -0.202, SE = 0.050, Z = 4.054, p < 0.001). The performance of the children in the subject-cleft condition was not significantly different from the adults' (Group$_{adults}$: b = -0.024, SE = 0.050, Z = -0.485, p = 0.628). Moreover, the differences of the proportion of looks to the ROI between the subject-cleft and subject-accented condition for the adult group were significantly larger than for the child group (Age Group$_{adults\ vs\ children}$ *Condition$_{subject-cleft\ vs\ subject-accented}$: b = 0.147, SE = 0.071, Z = 2.070, p < 0.05). However, the differences of the proportion of looks to the ROI between the subject-cleft and object-pseudocleft conditions were not significantly different for these two age groups (Age Group$_{adults\ vs\ children}$ *Condition$_{subject-cleft\ vs\ object-pseudocleft}$: b = 0.088, SE = 0.071, Z = 1.229, p = 0.222). To summarise, both child and adult participants looked more frequently to the alternative to the subject (the picture of the 'dog' in Figure 4) in the subject-cleft condition than in either the subject-accented or object-pseudocleft condition.

Figure 5. Proportion of looks to the alternative subject ('the dog' in Figure 4) against the mentioned subject ('the bird' in Figure 3) until the end of the trial with one standard deviation. The vertical dash line shows the end of the audio stimulus.
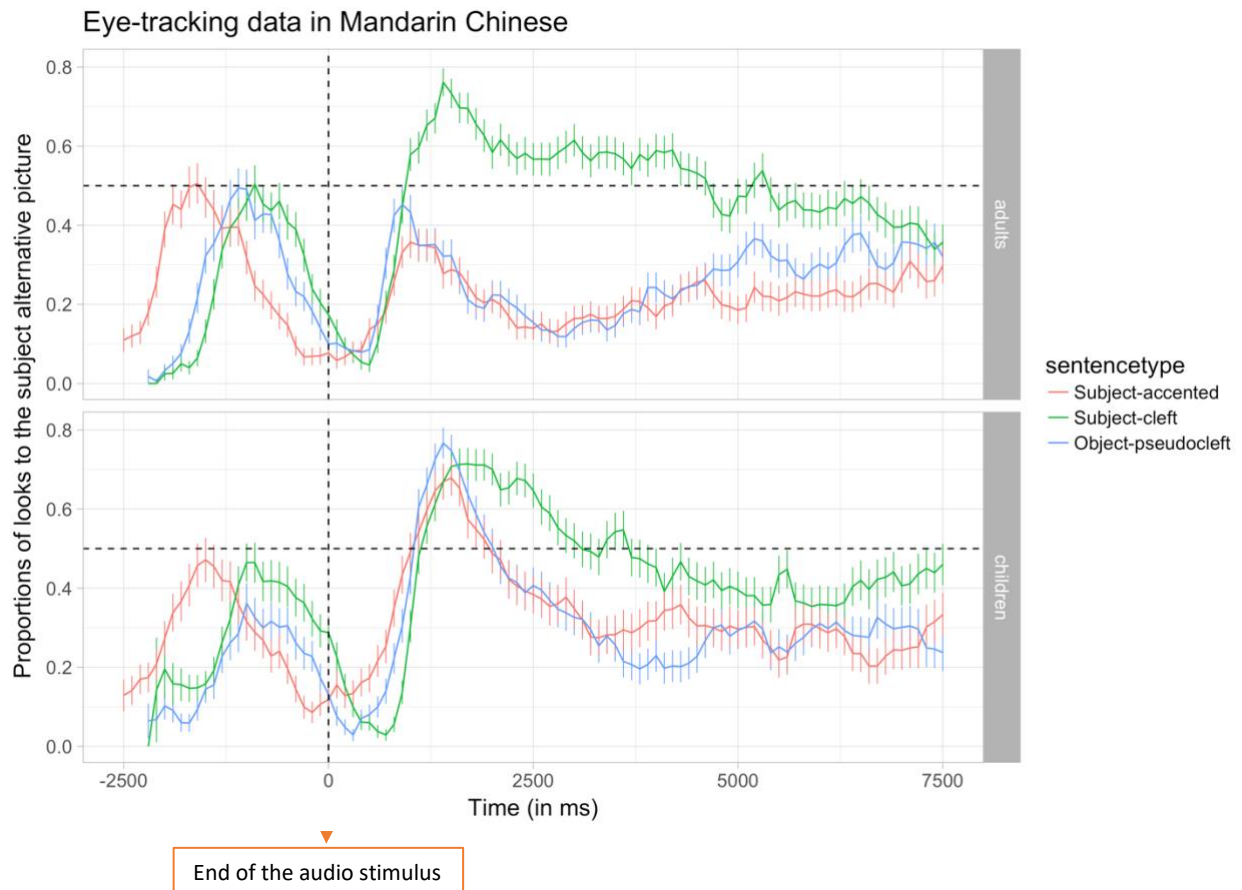
3.3 Study with German speakers

3.3.1. Method

3.3.1.1. Participants

53 monolingual, German-speaking five-year-old children (age range: 5.0–5.7, M = 5.2), reported as being typically developing children by their parents, participated in our study. Each was randomly assigned to one of the experimental conditions: 15 to the subject-accented condition, 13 to the object-accented condition, 13 to the subject-cleft condition, and 12 to the object-cleft condition. In addition, 59 German-speaking adult (under)graduate students at the University of Potsdam served as the control group. Each was also randomly assigned to one of the experimental conditions: 15 to the subject-accented condition, 14 to the object-accented condition, 15 to the subject-cleft condition, and 15 to the object-cleft condition. All participants were tested in the Baby Lab at the University of Potsdam, and had no visual and audio impairments at the time of testing.

3.3.1.2. Materials and design

Sentence-picture verification task with the VWP. The experiment consisted of 23 trials. These included an introduction trial, two practice trials, four filler trials, eight test trials, and eight control trials. In the test trials, the test sentences were false descriptions of the pictures, whereas the test sentences in the control trials were true descriptions of the pictures so as to balance the number of true and false statements. The trials were pseudo randomised so that no two same trial types would occur consecutively, and two different

order lists were created. The trials used the same pictures and design as those in the Mandarin trials. The only difference between these two experiments were the experimental conditions. All audio stimuli were pre-recorded by a female German native speaker. The average length of the audio duration in each condition was the following: 3.470s in the subject-accented condition, 3.766s in the subject-cleft condition, 3.457s in the object-accented condition, and 3.718s in the object-cleft condition. To run the experiment on the eye-tracker, all images were combined with the corresponding audio stimulus into video clips.

We included two prosodically-marked focus conditions and two syntactically-marked focus conditions, namely the subject-accented condition (Examples 7&8 in Table 2), the object-accented condition (Examples 9&10), the subject-cleft condition (Examples 11&12), and the object-cleft condition (Examples 13&14). In contrast to the Mandarin study, the object-cleft condition, instead of the object-pseudocleft condition, was investigated such that the structure is comparable with the subject-cleft sentences. The focus of the test sentence in the subject-accented and subject-cleft conditions was on the subject, whereas the focus of the test sentence in the object-accented and object-cleft conditions was on the object.

Table 2. Examples of the target and control trials for the four conditions

Target Trial	Control Trial
Example 7: Subject-accented condition Der VOGEL$_F$ hat die Flasche, stimmt das? The BIRDY$_F$ has the bottle, right it? 'The BIRDY$_F$ has the bottle, is that right?' Focus-congruent response: Nein, der Hund hat die Flasche. 'No, the dog has the bottle.'	**Example 8: Subject-accented condition** Der VOGEL$_F$ hat den Hammer, stimmt das? The BIRDY$_F$ has the hammer, right it? 'The BIRDY$_F$ has the hammer, is that right?' Focus-congruent response: Ja. 'Yes.'
Example 9: Subject-cleft condition Es ist der VOGEL$_F$, der die Flasche hat, It is the BIRDY$_F$, who the bottle has, stimmt das? right it? 'It is the BIRDY$_F$ who has the bottle, is that right?' Focus-congruent response: Nein, Es ist der Hund, der die Flasche hat. 'No, it's the dog who has the bottle.'	**Example 10: Subject-cleft condition** Es ist der VOGEL$_F$, der den Hammer hat, It is the BIRDY$_F$, who the hammer has, stimmt das? right it? 'It is the BIRDY$_F$ who has the hammer, is that right?' Focus-congruent response: Ja. 'Yes.'
Example 11: Object-accented condition Der Vogel hat die FLASCHE$_F$, stimmt das? The birdy has the BOTTLE$_F$, right it? 'The birdy has the BOTTLE$_F$, is that right?' Focus-congruent response: Nein, der Vogel hat den Hammer.	**Example 12: Object-accented condition** Der Vogel hat den HAMMER$_F$, stimmt das? The birdy has the HAMMER$_F$, right it? 'The birdy has the HAMMER$_F$, is that right?' Focus-congruent response: Ja.

'No, what the bird has is the hammer.'	'Yes.'
Example 13: Object-cleft condition	**Example 14: Object-cleft condition**
Es ist die FLASCHE$_F$, die der Vogel hat,	Es ist der HAMMER$_F$, den der Vogel hat,
It is the BOTTLE$_F$, that the birdy has,	It is the HAMMER$_F$, that the birdy has,
stimmt das?	stimmt das?
right it?	right it?
'It is the BOTTLE$_F$ that the birdy has, is that right?'	'It is the HAMMER$_F$ that the birdy has, is that right?'
Focus-congruent response:	Focus-congruent response:
Nein, es ist der Hammer den der Vogel hat.	Ja.
'No, it is the hammer that the bird has.'	'Yes.'

3.3.1.3. Procedures

We collected the verbal responses and eye gaze data. The data collection process and stimulus presentation was conducted with ClearView (version 2.5.1) on a Tobii 1750 binocular corneal reflection eye-tracker. The system tracks the eye gaze positions every 20ms with a tracking rate of 50 hertz. We followed the same procedures as with the Mandarin study.

3.3.1.4. Scoring of the response data

We applied the same scoring criteria as in the study with the Mandarin speakers: the subject corrections in the subject-accented and subject-cleft conditions were coded as

focus-congruent responses, and scored as correct; the object corrections in the subject-accented and subject-cleft condition were coded as focus-incongruent responses, and scored as incorrect. The object corrections that were provided in the object-accented and object-cleft conditions were coded focus-congruent responses and scored as correct, while the subject corrections in these two conditions were coded focus-incongruent responses and scored as incorrect. Moreover, if the participants only mentioned the target animals or items, the responses were still considered valid. For example, a response of a subject correction 'the birdy' in the subject-accented and subject-cleft conditions would be considered as valid and correct, whereas the same response in the object-accented and object-cleft conditions would be considered as valid but incorrect, and vice versa for the responses of the object corrections. Furthermore, there were instances in which the participants provided subject and object corrections in one response, such as 'The birdy has the hammer and the dog has the bottle', which we coded as invalid responses and excluded from the statistical analysis. In total, there were 28 invalid cases (5% of the data) for the adult participants (4 in the subject-accented condition, 16 in the subject-cleft condition, and 8 in the object-accented condition), and 37 invalid cases (8% of the data) in the child group (6 in the subject-accented conditions, 9 in the subject-cleft condition, 20 in the object-accented condition, and 2 in the object-cleft condition).

3.3.2. Results

3.3.2.1. Verbal responses

All participants were correct in the filler and control trials, meaning that none had to be excluded from the inferential statistics. A linear mixed-effect model (Baayen et al., 2008)

in R (R Core Team, 2016) was used in the R-studio environment (version 1.1.383). We sought to assess the impact of Sentence Conditions and Age Groups, and their interactions on the number of focus-congruent responses in the test trials. Further to these fixed factors, two random effects were included in the model, in which we considered by-subjects and by-items variance to the responses (the random slope) [16]. In order to compare their performance with adults, we established the contrast for Age Groups such that children was considered the 'baseline'. We tested whether the number of focus-congruent responses of the child group in the subject-accented condition was significantly different from those in the object-accented condition, whether the number of focus-congruent responses in the subject-cleft condition was significantly different from those in the subject-accented condition, and whether the number of focus-congruent responses in the object-cleft condition was significantly different from those in the subject-cleft condition. As for the interaction (Age Group*Condition), we examined whether the differences between the subject-accented and object-accented conditions, the subject-cleft condition and subject-accented conditions, and the object-cleft and subject-cleft conditions were significantly distinct between children and adults. Finally, we used the Wilcoxon test for the chance level tests with both participants in all conditions.

The results are summarised in Figure 6. There were no condition effects in the children's results (Condition$_{\text{subject-accented vs object-accented}}$: b = 3.587, SE = 2.145, Z = 1.672, p = 0.094; Condition$_{\text{subject-cleft vs subject-accented}}$: b = -2.136, SE = 2.031, Z = -1.502, p = 0.292; Condition$_{\text{object-cleft vs subject-cleft}}$: b = -3.497, SE = 2.132, Z = -1.640, p = 0.100). Moreover, due to the lower performance of the child group in the object-accented condition, the difference in the number of focus-congruent responses between the subject-accented and object-accented conditions was smaller for the adult group (Age Group$_{\text{adults vs children}}$

---

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

*Condition$_{\text{subject-accented vs object-accented}}$: b = - 5.798, SE = 2.829, Z = -2.050, p < 0.05). However, the differences in the number of focus-congruent responses between the subject-cleft and subject-accented conditions were statistically comparable across the age groups (Age Group$_{\text{adults vs children}}$*Condition$_{\text{subject-cleft vs subject-accented}}$: b = -0.614, SE = 2.823, Z = -0.218, p = 0.828), as were the differences between the object-cleft and subject-cleft conditions (Age Group$_{\text{adults vs children}}$*Condition$_{\text{object-cleft vs subject-cleft}}$: b = 3.886, SE = 2.884, Z = 1.347, p = 0.178).

Additionally, children performed at the chance level in the subject-accented (p = 0.19), the object-accented (p = 0.178), and the subject-cleft condition (p = 0.721), but performed below in the object-cleft condition (p = 0.009). The adult participants were above the chance level both in the subject-accented (p = 0.012) and object-accented condition (p = 0.003) but performed at the chance level in the subject-cleft (p = 0.580) and object-cleft conditions (p = 0.356).

Figure 6. Percentage of focus-congruent responses across Sentence Conditions and Age Groups. Each error bar shows the 95% confidence intervals for the mean
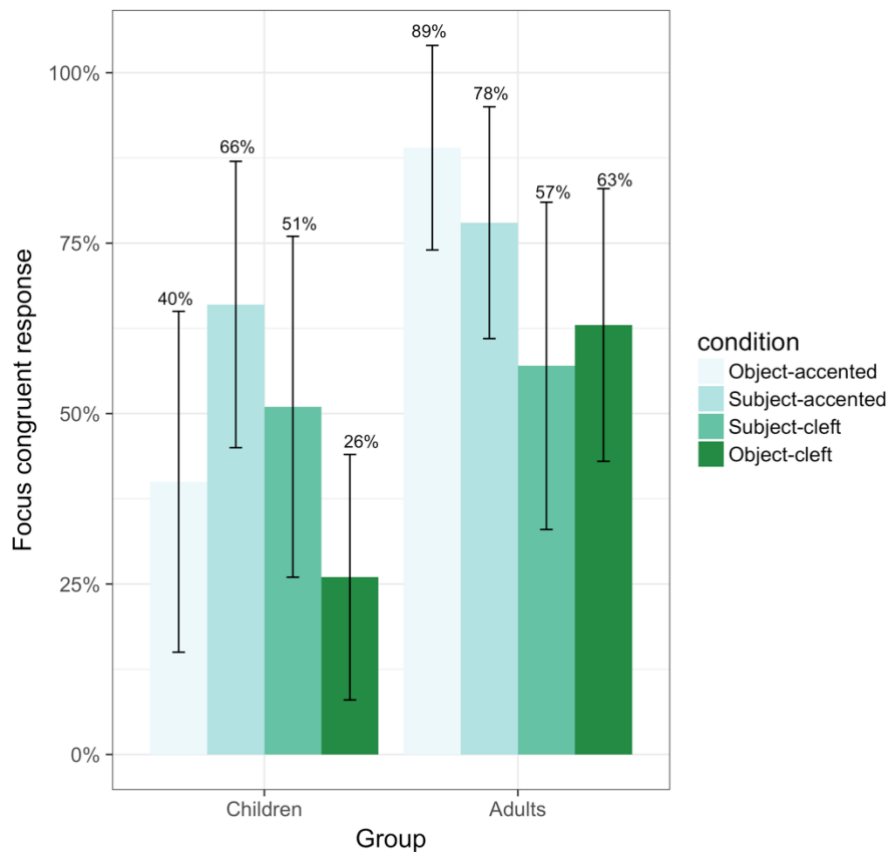
3.3.2.2. Looking pattern

3.3.2.2.1. Data treatment

The eye gaze data were processed using ClearView (version 2.5.1). For the test trials, we included only the looks towards the ROI which occurred after the end of the audio stimuli for inferential statistics. As before, our purpose was to assess how the prosodic and syntactic information guided the participants' attention. Trials were excluded in

accordance with the 50% track loss criterion (6.8% for the adult data and 10.3% for the child data).

3.3.2.2.2. Region of Interest

The ROI was defined in the same manner as in the Mandarin study (see Figure 4). For the subject-accented and subject-cleft conditions, the looks to the ROI were considered as the activation of the alternatives for the sentence subject. In contrast, the ROI in the object-accented and object-cleft conditions did not contain the alternatives for the focused object. Therefore, we expected fewer looks to this ROI in these conditions, and data from these two conditions served as a contrast dataset for the other two conditions.

3.3.2.2.3. Data analysis

A linear mixed model (Baayen et al., 2008) in R (R Core Team, 2016) under R Studio (version: 1.1.383) was used in order to assess the impact of Sentence Conditions and Age Groups, and their interactions, on the proportions of looks to the ROI. We included an additional three random effects in which by-subjects, by-items, and by-position variance to the proportions of looks to the ROI were considered[17]. Again, to be comparable to the model for the response data (see Section 3.3.2.2.), the contrast for Age Groups set children as the 'baseline' so as to compare their performance with adults. First, for the child group, we tested whether the proportion of looks to the ROI in the subject-accented condition were significantly different from those in the object-accented condition. This allowed us to gain an insight into which of the above two conditions the participants preferred

---

[17] Proportions of looks to the ROI~group*condition+(1|id)+(1|item)+(1|position)

looking to as alternative subjects. Second, we examined whether the proportions of looks to the ROI in the subject-cleft condition were significantly different from those in the subject-accented condition. We did so as this would allow us to ascertain whether the participants treated the alternative subjects in the same manner in the prosodically-marked focus (the subject-accented condition) and in the syntactically-marked focus (the subject-cleft condition). Finally, we analysed whether the proportions of looks to the ROI in the object-cleft condition significantly differed from those in the subject-cleft condition with the intention of understanding the participants' condition preference in terms of alternative subjects. As for the interaction (Age Group*Condition), we examined whether the differences between the subject-accented and object-accented conditions, the subject-cleft and subject-accented conditions, and the object-cleft and subject-cleft conditions were significantly distinct between the child and adult group.

The results can be seen in Figure 7. Regarding the eye-tracking results, there were no Sentence Condition effects for the child group (Condition$_{\text{subject-accented vs object-accented}}$: b = 0.080, SE = 0.063, Z = 1.278, p = 0.203; Condition$_{\text{subject-cleft vs subject-accented}}$: b = -0.071, SE = 0.063, Z = -1.140, p = 0.256; Condition$_{\text{object-cleft vs subject-cleft}}$: b = 0.072, SE = 0.066, Z = 1.084, p = 0.280). Moreover, due to the higher proportion of looks to the alternative subject of the child group in the object-accented condition, the differences of the proportion of looks to the ROI between the subject-accented and object-accented condition were larger for the adult group (Age Group$_{\text{adults vs children}}$*Condition$_{\text{subject-accented vs object-accented}}$: b = 0.374, SE = 0.084, Z = 4.435, p < 0.001). However, there were no differences in the proportion of looks to the ROI between the subject-accented and subject-cleft condition, or between the object-cleft and subject-cleft condition for both age groups (Age Group$_{\text{adults vs children}}$*Condition$_{\text{subject-cleft vs subject-accented}}$: b = -0.155, SE = 0.083, Z = -1.866, p = 0.065; Age Group$_{\text{adults vs children}}$*Condition$_{\text{object-cleft vs subject-cleft}}$: b = -0.071, SE = 0.085, Z = -0.835, p =

0.405). However, one drawback of this model is that the way in which the adults treated the alternative subject in the conditions could not be represented. Therefore, we applied the same model again but set adults as the baseline[18]. Regarding the eye-tracking data of the adult group, the proportion of looks to the alternative subject in the subject-accented condition were significantly higher than in the object-accented condition (Condition$_{\text{subject-accented vs object-accented}}$: b = 0.456, SE = 0.063, Z = 7.204, p < 0.001), while the proportion of looks to the alternative subject in the subject-cleft condition were significantly lower than in the subject-accented condition (Condition$_{\text{subject-cleft vs subject-accented}}$: b = -0.228, SE = 0.061, Z = -3.710, p < 0.001). Moreover, we found no difference in the proportion of looks to the alternative subject between the subject-cleft and object-cleft conditions (Condition$_{\text{object-cleft vs subject-cleft}}$: b = -0.008, SE = 0.060, Z = -0.134, p = 0.893).

---

[18] The model is the same as shown in footnote 17: ROI~group*condition+(1|id)+(1|item)+(1|position)

Figure 7. Proportion of looks to the alternative subject ('the dog' in Figure 4.) until the end of the trial with one standard deviation. The first vertical dash line shows the end of the audio stimulus
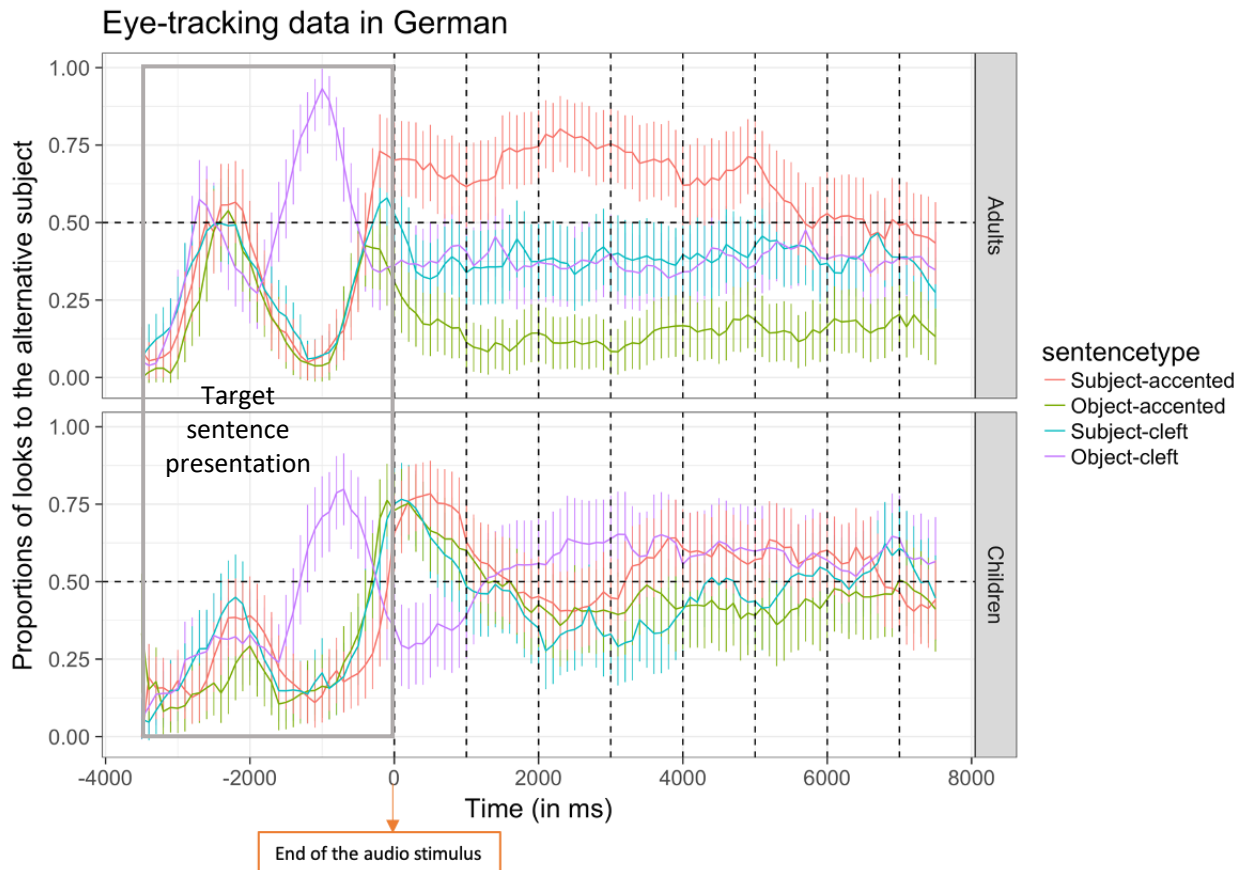
3.4 Discussion

In the current study, we have scrutinised whether speakers of Mandarin (a tone language) and speakers of German (an intonation language) could understand prosodically-marked or syntactically-marked focus by using a sentence-picture verification task with the visual world paradigm. We employed a sentence-picture verification task and the eye-tracking method to observe developmental performance of understanding prosodic and syntactic information on focus marking. Additionally, eye-tracking allowed us to observe more closely how different focus structures guide the participants' attention. The results

revealed two main findings. First, Mandarin speakers (both children and adults) more effectively exploited syntactic than prosodic cues of focus marking, whereas prosodic information (compared to syntactic) played a more important role in understanding focus in sentences for German speakers. Cross-linguistically, it has been demonstrated that children follow language specific cues in their ambient language from an early stage (in this case, from the age of five). We shall first discuss the details of the Mandarin results.

The results of the sentence-verification task showed that both Mandarin-speaking children and adults performed below the chance level in the subject-accented condition. That is, not only adults but also children, instead of following the prosodic information, checked word order information and hence corrected the sentence objects. This suggests that Mandarin speakers emphasise word order information in their focus interpretation. Moreover, our results align with the study reported in Chapter 2, as the objects (in the default focus position) were corrected most often in the subject-accented condition by Mandarin-speaking adults (88%) and children (80%). Again, these results provide evidence that Mandarin, as a tone and topic-prominent language, is quite restricted in the use of prosodic information despite prosodically-marked subject focus being available within it (Shih, 19898; Xu, 1999). This finding is not compatible with Chen's (1998) finding that word-order information (new information before given information) is more significant to apprehending focus in sentences than prosodic information for Mandarin-speaking adults, but not for children. Secondly, Mandarin speakers (adults and children) could mostly identify focus by exploiting the specific syntactic structures: the objects were most often corrected in the object-pseudocleft condition by adults (88%) and children (79%) while, in the subject-cleft construction condition, the subjects were often corrected by adults (70%), but more infrequently by children (39%). This confirms the hypothesis that the object-pseudocleft construction is easier to comprehend than the subject-cleft

construction for Mandarin-speaking children, since the syntactically-marked focus is aligned with the default focus position in the former. This assumption aligns with Chen's (1998) finding that children (across different age groups) performed more ably in the object-pseudocleft construction condition than in the subject-cleft construction condition.

However, although the response data showed that children in the subject-cleft condition performed at the chance level, and the adults above it, the eye-tracking data reveals a different picture. Both Mandarin-speaking adults and children showed a higher proportion of looks to the alternative subject in the subject-cleft condition than in the subject-accented and object-pseudocleft conditions. This suggests that the five-year-old children processed the subject-cleft sentences in an adult-like fashion and assigned focus to the sentence subject. It is worth noting that this is not the first study to present discrepancies between verification and eye-tracking results (Höhle et al., 2016; Zhou et al., 2012). Höhle et al. (2016) found that four-year-old German-speaking children processed the pre-subject only sentences in the same way as adults, despite the child participants performing below the chance level in the pre-subject only condition. In Zhou and his colleagues' (2012) study, Mandarin-speaking children and adults were asked to verify pre-subject only sentences with an accent either on the head-noun or the subject modifiers, combined with the VWP. In the judgement results, the adult participants linked the focus particle 'only' to the head-noun or the modifier following the prosodic information. In contrast, the child participants applied a default strategy by only correcting the modifiers regardless of the positioning of the prosodic information. However, in the eye-movement data, the child looking patterns mirrored those of the adults, yet children were slower to fixate on the targets. Accordingly, Zhou and his colleagues assumed that, while children can process prosodic information, they can make little use of this information in their final judgments about a sentence's correctness due

to their still-limited processing faculties. They further proposed that the representation of different levels of linguistic properties – in this case, the mapping of syntax and phonology – is challenging for children, but not for adults. Therefore, the eye-tracking results (implicit responses) facilitate our understanding as to how Mandarin-speaking children comprehend the subject-cleft construction. For instance, while Mandarin-speaking children looked more frequently to the subject alternatives in the subject-cleft condition, they made their final verification decisions on correcting objects – quite possibly due to Mandarin's topic-prominence property. Following Zhou and his colleagues (2012), we assume that the child participants failed in representing the information they had collected according to the eye-tracking pattern because of their limited processing recourses, meaning that said information no longer guided their verification responses. Overall, the results confirm our hypothesis that the object-pseudocleft sentences are easier to understand than the subject-cleft sentences for Mandarin-speaking children given that the objects were focus-marked by two means – the occurrence of the object in the default focus position at the end of the sentence and the marked syntactic structure of the pseudocleft construction.

The response data and eye-tracking results showed that, contrary to Mandarin speakers, German adults performed better in the prosodically-marked focus conditions (the subject-accented and object-accented conditions) than in the syntactically-marked focus conditions (the subject-cleft and object-cleft conditions). From the response data, we can see that the adults performed above the chance level in the subject- and object-accented conditions, but only performed at the chance level in the subject- and object-cleft conditions. These findings are consistent with Szendrői et al. (2017), who found that German speakers were sensitive to prosodic information of focus marking. However, the results of the German-speaking children were less clear cut. In view of the results in the

prosodically-marked focus condition, their performance in the subject-accented condition was comparable to the adults' performance, although they performed at the chance level performance in the object-accented condition regarding the focus-congruent responses. However, Szendrői and her colleagues (2017) found that the German children performed more highly in the object-accented condition than in the subject-accented condition. As yet, there is no clear explanation for such an outcome. One potential reason could be the application of a between-subject design. It is possible that the German-speaking children who participated in the subject-accented condition were, in general, more able to exploit prosodic information to identify focus than the children in the object-accented condition. However, there were no differences of gender or average age between the two groups (average age in subject-accented condition: 5.03, average age in object-accented condition: 5.02; female number in subject-accented: 8; male number in subject-accented: 7; female number in object-accented: 7; male number in object-accented: 6). Furthermore, it could be that the additional prosodic information confused the child participants. Cross-linguistically, and certainly true for German, the default focus position comes as the end of the mostan embedded clause. Indeed, the default focus in the object-accented condition should be the object, even without additional accents on the object (Xu, 2004; Reinhart, 2004). Therefore, one would expect a generally high performance in object-focus sentence. On the other hand, Müller and her colleagues Höhle (2013) assumed that, because of the conflict of the information structure in the pre-subject only sentences, German-speaking children would face comprehension difficulties. This conflict also holds for the subject-accented sentences in the current experiment. However, the German children still performed more effectively in the subject-accented, rather than object-accented, sentences. In the response data, the double corrections 'the birdy has the bottle and the dog has the hammer' occurred more frequently in the object-accented

condition than in the other three. These results were comparable with the six-year-old German-speaking children in Szendrői et al.'s (2017) study, in which there was a large number of double corrections (22% of the time), and an almost equal number of subject corrections and object corrections in the object-accented condition (40% vs 38%). The higher numbers of double corrections, and the almost equal number of the subject and object corrections in the object-accented condition in both studies indicates that German-speaking children have difficulties in identifying focus when additional prosodic information on the object is given.

The eye-tracking results further confirmed that German speakers were sensitive to the prosodically-marked focus, but not to the syntactically-marked focus. First, both the adults and children looked more frequently to the subject alternative (e.g., the dog) in the subject-accented condition, but not in the object-accented condition. However, unlike adults, children's looks to the subject alternative in the object-accented condition did not especially differ from their looks to the subject-alternative in the subject-accented condition. This result aligns with those of the sentence picture verification task that children corrected the subject while visually checking the subject alternative in the object-accented condition. Regarding the adults' eye-tracking results, we observed no different looking patterns between the subject- and object-cleft conditions, thereby indicating that the adults showed no preference for looking at the subject alternative (e.g., the dog) compared to the mentioned subject referent (e.g., the bird). Again, this aligns with the response data that the adults' performance in the sentence-verification task was not different from the chance level in these two conditions.

Both the response and eye gaze data suggest that the cleft construction might not be valid for indicating focus in German. The response data showed that even the performance of German-speaking adults did not differ from the chance level in the cleft

conditions (the focus-congruent responses for the subject-cleft were 57% and 63% for the object-cleft). This suggests that they did not follow the syntactic cues provided in the cleft construction to either establish the subject- or object-alternative set. As for the German child participants, their performance in the subject-cleft condition was comparable with the adult participants in that they performed at the chance level (51%). However, in the object-cleft condition, the child participants performed below the chance level (26% focus-congruent responses), suggesting that they might follow a specific response strategy in this condition by systematically correcting the subject. In German, information structure of the left periphery of a sentence is considered the topic rather than the focus (Musan, 2010). This has two implications for the current study. First, it might be that the child participants counted the phrase '*Es ist die Flasche* [It is the bottle]' as the topic. Following Reinhart (1982), topic indicates what the sentence is about, and it is not necessary being the sentence subject. Therefore, information concerning the bottle (topic) became relevant for them to answer the question: on the stimulus, the dog has the bottle. Second, it is also valid in German that the focused element falls in the end of the sentence, i.e., the default focus position. The children considered the relative clause '*die der Vogel hat* [that the bird has]' as the focus, and so they looked for the alternative subject and corrected it (i.e., the dog). As the matter of fact, word order information is also significant for German-speaking children while identifying focus in sentences. Overall, the result pattern of the German children corresponded to Lampert and Kinsborne's (1980) findings on English-speaking children: the object-cleft was more difficult to understand than the subject-cleft[19].

Our empirical findings concerning the German cleft constructions provide support to what has previously been proposed in the literature. Following Prince (1978),

---

[19] To the best of our knowledge, there was no empirical evidence in German concerning acquisition of the cleft construction, meaning our study would be first to provide it.

Lambrecht (1994) assumed that the cleft construction could be further divided into three subtypes concerning its information structure: predicate-focus, argument-focus, and sentence-focus structure[20]. In a predicate-focus structure the clefted constituent marks the given information followed by the new information whereas, in an argument-focus structure, the new information is denoted by the clefted constituent and what follows it is the given information (Examples a and b). In a sentence-focus structure, the whole sentence represents the new information, i.e., the clefted constituent as well as the predicate (Example c).

a. Predicate-focus structure

   Context: John ate the chocolate and the cake?

   It was the CHOCOLATE that Kevin ate. (Given-New)

b. Argument-focus structure

   Context: Who ate the chocolate?

   It was JOHN who ate the chocolate. (New-Given)

c. Sentence-focus structure

   Context: What happened?

   It was yesterday that John left us. (New)

Interestingly, when listeners comprehend the cleft construction, the clefted constituent could be either represented as given or new information – an assumption which corresponds to our findings in both the response and eye-tracking-data with German-speaking adults. In the response data, the performance of German-speaking adults was at the chance level. In the eye-tracking data, the proportion of looks to the alternative subject and the proportion of looks to the sentence subject were the same for German-speaking adults. That is, they did not distinguish the information of the clefted constituents of 'new'

---

[20] In the book, Lambrecht mainly discussed English and French cleft construction sentences.

or 'given', but instead just decided randomly. Secondly, our findings also relate to Dufter's (2009) observations in his corpus study. In his corpus data, the total number of the cleft construction sentences in German were the lowest when compared to other European languages (i.e., Spanish, Portuguese, Italian, and French). He argued that the less restricted ways of focus marking in German – not only through clefting, but also by adding focus particles or varying word order – might cause the low frequency of the cleft construction. Consequently, this low occurrence may explain the findings concerning the understanding, and inconsistent interpretation, of the cleft construction. The current study is the first empirical research with German-speaking adults and children that investigates the comprehension and acquisition of the cleft construction in German. Further research may help us more comprehensively understand how speakers of German use and understand this construction.

Let us return to the German and Mandarin eye-tracking results for a closer inspection. Both Mandarin-speaking children and adults, and German-speaking adults, more frequently looked to the subject alternatives in the conditions where the specific cues were available for the corresponding languages: more looks to the alternative subjects showed up in the subject-cleft condition for Mandarin speakers (adults and children) whereas only German-speaking adults looked more to the alternative subjects in the subject-accented condition. However, both Mandarin- and German-speaking children displayed only a chance level performance in their verbal responses in the subject-cleft and subject-accented conditions. It seems that the Mandarin child parsers would direct their attention to the target picture they were supposed to look at, but they could not apply the results of this processing in the explicit task (Brandt-Kobele and Höhle, 2010; Zhou et al., 2012; Höhle et al., 2016). This suggests that the eye-tracking method can be useful for providing a detailed picture of children's language processing. That is,

while children may be unable to conduct an explicit task by providing the correct responses, but the eye-tracking data suggest that the focus structure of the sentences has an effect on their interpretation process. Considering only the response data might have led us to the erroneous conclusion that Mandarin-speaking children could not follow the syntactic cues of focus marking. However, through using the eye-tracking method, we were able to gain valuable insights into how the specific language cues guided the participants' attention, as well as directly compare the eye looking pattern of adults and children.

One fact ought be mentioned in the following. In the present research, a between-subject design was used which may have led to different variance from condition to condition in the results. A pilot study (using a within-subject design) with Mandarin adults was conducted beforehand, and we observed that the task was surprisingly demanding even for the adult participants – indeed, the number of focus-congruent responses in the subject-cleft condition was very low (26%). Therefore, it is appropriate to apply the present scheme (i.e., a between-subject design) to investigate the research questions.

The current study fulfils the needs of cross-linguistic research on languages with relevant structural differences, in this case German as an intonation language and Mandarin as a tone language. One of the most important cross-linguistic implications here is that the syntactic cues of focus marking are more crucial than the prosodic cues for Mandarin speakers, while the reverse is true for German speakers. The current results are in line with Szendrői and her colleagues (2017) and Chen et al. (2019), in that they confirm that children are attuned to the language specific cues from an early age. In addition, eye-tracking seems to be a highly sensitive method for investigating children's sentence comprehension and it helped us gain a more in-depth understanding of how

Mandarin children understand focus structures. Another significant finding of our research is that the cleft construction seems to be valid for marking focus in Mandarin, but not in German. This not only supports the fact that Mandarin is a topic-prominent language, but also shows that focus can be marked in German in a multitude of ways. For instance, in Mandarin, the first element of the sentence serves as the topic, meaning that a specific construction (i.e., the cleft construction) is needed if the first element of a sentence is a focus rather than a topic. In German, however, adding focus particles, changing the word order, or altering the accentuation also serve to mark focus. In order to shed more light on understanding the syntactically-marked focus in both Mandarin and German, further research should seek to examine a child corpus in which the frequency use of the cleft construction from the mother input can, along with production data, be presented. Moreover, regarding our German results, it seems that there are more means to mark focus in German than in Mandarin. Accordingly, the results from the German child participants were not as clear as those relating to their Mandarin counterparts, thereby necessitating further research into how German-speaking children acquire these different means (in the current study, we explored the prosodically-marked focus and the cleft construction for marking focus).

4 Understanding the Focus Particle '*Zhiyou*' ('Only') in Mandarin—Evidence from Five-Year-Olds

Much research has been done cross-linguistically to understand how speakers utilise different ways of marking focus, such as changing word order, adding prosodic information, using cleft sentences and so on. However, there is a dearth of research that has investigated how children acquire these various means. In the previous chapters, we have explored how Mandarin and German speakers, both adults and children, apply prosodic and syntactic information to identify the focus in sentences. The results suggest that Mandarin-speaking adults and children could better exploit syntactic cues (i.e., word order and the cleft construction) than prosodic cues (i.e., focus accent) as a means of focus marking, whereas German-speaking adults and children are more sensitive to prosodic information for identifying the focus in sentences. In the current chapter, we will investigate another focus structure marked by the focus particle 'only', and we aim to discern whether five-year-old Mandarin-speaking children can understand this type of structure. Additionally, we will examine whether prosodic information can help Mandarin speakers to resolve the ambiguity in sentences that contain the focus particle 'only'.

4.1 Introduction

Cross-linguistic research has shown that children up to school-going age struggle to comprehend sentences that contain the focus particle 'only' (English: Crain et al., Ni and Conway, 1994; Paterson et al., Liversedge, Rowland and Filik, 2003. German: Müller et al.,

Schulz and Höhle, 2011. Mandarin: Yang, 2002; Zhou et al., Su, Crain, Gao and Zhan, 2012. Portuguese: Costa and Szendrői, 2006).

Consider Example 1-3:

1. Pre-object-only sentence:

   John only ate a [banana]$_F$[21].

   (John only ate a banana, and nothing else was present in the discourse model.)

2. Pre-VP-only sentence:

   John only [ate a banana]$_F$.

   (John was only eating a banana, and no other activity was present in the discourse model.)

3. Pre-subject-only sentence:

   Only [John]$_F$ ate a banana.

   (Only John ate a banana, and no one else was present in the discourse model who ate a banana.)

Following Rooth (1992), focus indicates the presence of alternatives in the current discourse model. In Example 1, the scope of the focus particle 'only' is over the object *banana*, and it marks the object as the focus. The focus particle 'only' is an exclusive particle, excluding the referent of the focus element from the set of alternatives. Thus, Example 1 emphasises that what John eats is a banana, and no other object that is contained in the set of alternatives is present in the discourse model. Contrary to Example 1, the focus in Example 2 is the verb phrase: the scope of the focus particle 'only' is over the verb phrase *ate a banana*, and this stresses that what John did was eat a banana, and he performed no other action that was contained in the set of alternatives present in the discourse model. In Example 3, the subject *John* is the focus. Accordingly, this sentence

---

[21] See Footnote 1.

indicates that John is the only one who eats a banana, and there is no one else in the current discourse model who eats a banana. Hereafter, sentences with an object focus will be called **pre-object 'only' sentences,** sentences with a verb phrase focus will be called **pre-VP 'only' sentences** and sentences with a subject focus will be called **pre-subject 'only' sentences**, irrespective of the exact position of the focus particle.

Previous research has indicated that three- to four-year-old children can understand pre-object 'only' sentences but are unable to interpret pre-subject 'only' sentences correctly (Berger and Höhle, 2012; Crain et al., 1994; Müller et al., 2011; Yang, 2002). To be able to access the meaning of sentences with the focus particle 'only', the parser needs to integrate several different linguistic levels (e.g., syntax, prosody, lexicon, pragmatics). Take the sentence *Ellen only gave John a banana* as an example. To interpret the sentence, the lexical meaning of the focus particle 'only' and the domain of focus *gave John a banana* need to be identified. Meanwhile, the corresponding sets of alternatives should be established, and depending on the accent placement, different sets of alternatives are possible. When the accent is placed on the indirect object *John*, the set of alternatives would be other persons (e.g., *Mary*, *Paul* or *Bill*), whereas when the direct object *banana* is accented, the set of alternatives would be other things (e.g., *apple*, *peach* or *watermelon*).

So far, there is no consensus among language acquisition researchers regarding which linguistic level or levels cause the comprehension problem that children have with these constructions. In 1994, Crain and his colleagues reported on an earlier study that Crain, Philip, Drozd, Roeper and Matsuoka (1992) conducted with English-speaking children, where more than half of the three- to six-year-old participants (21 out of 38) interpreted the pre-subject 'only' sentence *Only the cat is holding the flag* (The cat but no other animal is holding the flag.) as having the meaning of the pre-VP 'only' sentence *The*

cat is only holding the flag. (The cat is holding the flag but not doing any other action.) Accordingly, Crain and his colleagues (1994) proposed that due to syntactic misrepresentation, children were unable to parse the pre-subject 'only' sentences correctly: the child participants assigned the scope of the focus particle 'only' to the verb phrase and did not restrict it to the subject (the syntactic proposal). Similarly, Notley and her colleagues (2009) argued that children could not understand pre-subject 'only' sentences because they misanalysed the scope domain of the focus particle 'only'. In their study, they tested four- to five-year-old Mandarin-speaking children with pre-subject 'only' sentences by using the truth-value judgement task. Their results demonstrated that despite different test conditions, the focus particle was constantly associated with the object instead of the subject: in the true condition, the context was that Mr Pig got a gold coin and a silver coin and Mr Horse got a gold coin. The test sentence was *Only Mr Pig got a silver coin*. In the false condition, the context was that Mr Cat ate a fish and a carrot and Mr Rabbit ate a carrot. The test sentence was *Only Mr Cat ate a carrot*. In both conditions, when the Mandarin-speaking children rejected the statement, they persistently corrected the object instead of the subject. (i.e., *No, Mr Pig also got a gold coin* and *No, Mr Cat also ate a fish.*)

On the other hand, Paterson and his colleagues (2003) assumed that children misunderstood the pre-subject 'only' sentences because they had difficulties representing the set of alternatives, which is necessary to be able to comprehend the sentences with the focus particle 'only'. In their study, by using the forced-choice picture-selection task, they examined English-speaking adults and children aged four to twelve years not only with sentences with the focus particle 'only' but also with sentences without the focus particle. In this way, they checked whether the participants would submit different responses for sentences with and without the focus particle 'only'. If the participants

submitted the same responses for these two sentences, it showed that they ignored the information of the focus particle 'only' and did not construct a set of alternatives. Regarding their results, all participants comprehended the sentences without the focus particle 'only' easily. Moreover, the six-year-old and younger participants submitted the same responses for the with-'only' test sentences as for the without-'only' test sentences; therefore, the authors assumed that the child participants did not consider the set of alternatives while figuring out the meaning of the sentences with the focus particle 'only'. However, results from a later study by Paterson, Liversedge, White et al., Filik and Jaz (2006) indicated that the effect found in the previous study was due to the pragmatically infelicitous task.

Müller, Höhle and Schulz (2013) took another approach. In their study, three different types of sentences with the focus particle 'only' were included to examine four- to six-year-old German-speaking children by using the sentence–picture verification task. Two conditions—pre-subject 'only' and pre-object 'only'—were included in sentences with the canonical word order SVO, e.g., *Nur die Maus hat den Ballon* (Only the mouse has the balloon.) and *Die Maus hat nur den Ballon.* (The mouse only has the balloon.) In addition, a third condition contained pre-subject 'only' sentences in the non-canonical word order OVS, e.g., *Den Ballon hat nur die Maus.* (Literally: the balloon has only the mouse, which means the same as only the mouse has the balloon.) The results indicated that the performance of the German-speaking children for pre-subject 'only' sentences was the same between the SVO and OVS sentences. That is, German-speaking children interpreted both types of sentences (i.e., the canonical pre-subject 'only' sentences and the non-canonical pre-subject 'only' sentences) as pre-object 'only' sentences. If the German-speaking children had misanalysed the syntactic scope of the focus particle 'only', they would have shown better performance with the non-canonical pre-subject 'only'

sentence since the focus particle of the sentence had no scope over the object that appeared before the subject. Müller and her colleagues argued that the mismatch of information and syntactic structure led to the comprehension difficulties: in the pre-subject 'only' sentence, the subject was the sentence focus, which is in conflict with the subject's status as the default topic in German. It is then challenging for the children to interpret the meanings of the pre-subject 'only' sentences. Their results provided evidence against the syntactic proposal of Crain and his colleagues (1994), in which they assumed that children represented the pre-subject 'only' sentences as the pre-object 'only' sentences because they misrepresented the syntactic scope of the focus particle 'only'. (See above.)

Similarly, Yang (2002) proposed that the need for reanalysing the information structure status of the subject in pre-subject 'only' sentences resulted in Mandarin-speaking children struggling to comprehend these sentences: as in German, the subject is typically the topic in Mandarin and is, therefore, initially considered to be the topic (Li and Thompson, 19891); however, in a pre-subject 'only' sentence, the subject is the focus. In Yang's study, forty-eight four- to eight-year-old Mandarin-speaking children were tested by using the sentence–picture verification task, and again, the results showed that Mandarin-speaking children up to the age of six years interpreted the pre-subject 'only' sentences in the same way as they did the pre-object 'only' sentences. In addition, Mandarin-speaking children had adult-like performance only from eight years on: only at this age did more than half of the participants (eight out of the 12 children) represent the pre-subject 'only' sentences correctly.

Prosodic information can be used as a tool for marking focus (Selkirk, 1995; Gussenhoven, 2008), which was addressed in Chapter 2. In the current chapter, a further study regarding prosodic information that can be applied to resolving structure

ambiguities will be carried out. Cross-linguistic research has shown that prosodic information is supportive for adults to solve structural ambiguities, but this is not the case for children (Snedeker and Trueswell, 2001; Choi and Mazuka, 2003; Snedeker and Trueswell, 2003). However, a later study by Snedeker and Yuan (2008) found that English-speaking children, as well as English-speaking adults, could apply prosodic information to resolve structural ambiguities in sentences. In their study, English-speaking children (four to six years old) and adults were tested using the visual world paradigm, in which the participants were asked to listen to sentences such as *You can feel the frog with the feather*. This sentence has two interpretations: either the prepositional phrase is attached to the verb phrase *feel the frag* as an instrument, or it is taken as the modifier of the noun phrase *the frog*. The prosodic information was included to resolve the ambiguity: there was either an intonational phrase boundary after the verb (as in the modifier prosody condition) or after the direct object (as in the instrumental prosody condition). While the participants performed the action, their eye gaze data were collected. The action data, as well as the eye gaze data, indicated that English-speaking children and adults could apply the prosodic information. Further, the child participants demonstrated a 500ms delay while processing the prosodic information compared to the adult participants.

Let us turn to the literature on prosodic effects when resolving structural ambiguities in Mandarin. Zhou and colleagues, Su, Crain, Gao and Zhan (2012) investigated whether Mandarin speakers (adults and four-year-olds) could exploit prosodic information to resolve structural ambiguities and interpret speech acts. The study concerning the speech acts will not be discussed here, as this topic is beyond the scope of the current research. In the structural ambiguity study, the question of whether Mandarin-speaking children and adults could exploit prosodic information to resolve the

ambiguity in sentences with the focus particle 'only' was addressed by using the sentence–picture verification task with the visual world paradigm. A sentence like *Zhiyou Xiaoming de niaozhong shi huangsede* (Only Xiaoming's clock is yellow.) contains two possible interpretations: either Xiaoming's clock is yellow and no one else's clock is yellow in the context of the discourse, or Xiaoming's clock is yellow and nothing else is yellow in the context of the discourse. Adding an accent either on the modifier—*Only XIAOMING'S$_F$ clock is yellow*—or on the head noun—*Only Xiaoming's CLOCK$_F$ is yellow*—would determine the sentence's interpretation. In the first case, the accent on the modifier indicates that Xiaoming's clock is yellow and no one else's clock is yellow. In contrast, the sentence where the accent is on the head noun means that Xiaoming's clock is yellow and nothing else is yellow. As for the visual stimulus, two characters—a boy, Xiaoming, and a girl, Xiaohong—were presented, and Xiaoming was holding a yellow clock and green scissors, whereas Xiaohong was holding a yellow clock and yellow scissors. The verification results showed that children had an adult-like performance in the accent-on-the-modifier condition (modifier accented condition) but not in the accent-on-the-head-noun condition (head noun accented condition). In the head noun accented condition, children gave the same response as in the modifier accented condition by pointing out that not only did Xiaoming have a yellow clock but Xiaohong also had a yellow clock. Moreover, in the eye gaze data, children showed adult-like performance but with a 400ms delay for both conditions: both adult and child participants looked more to Xiaoming's clock in the modifier accented condition, while they looked more to the other yellow thing that the contrastive character, Xiaohong, possessed in the head noun accented condition. Overall, although Mandarin-speaking children could not provide adult-like responses, their looking patterns were similar to the adults'. To summarise, Zhou and his colleagues concluded that children seemed to have a default interpretation—the modifier

interpretation—regardless of the condition (head noun accented or modifier accented). Contrary to the adults, children preferred interpreting the sentences as having an accent on the modifier. Moreover, concerning the discrepancy between the verification results and the eye gaze results, Zhou and his colleagues assumed that children showed longer processing because the mapping of syntax and prosody was not yet automatic for them.

In Zhou and his colleagues' study, the children could seemingly apply the prosodic information to disambiguate the structural ambiguities but behaved differently from the adults: they had a default interpretation of a narrow focus on the modifier. However, in their study, they only tested pre-subject 'only' sentences. Zhou and his colleagues showed that prosodic information guided Mandarin children's focus processing in sentences with pre-verbal subjects but that they could not use this information in the explicit sentence judgement task. To my knowledge, this study is the first one in Mandarin concerning the use of prosodic information as a tool to resolve structural ambiguities in sentences with the focus particle 'only'. Based on the results, the question emerges regarding whether Mandarin-speaking children would show similar effects of prosodic information in a pre-subject 'only' sentence when the sentence includes an object in addition to the subject. On the other hand, Chen and her colleagues (2019) showed that Mandarin-speaking children preferred the default focus interpretation for the SOV sentences in a sentence verification task: the object was considered as the focus, even though the subject was accented. An interesting issue arose regarding whether this default interpretation could be counteracted by sentences with the additional prosodic information on the head noun or on the modifier of the subject as the focus element. Therefore, in the current study, we extend the cross-linguistic approach in this research area by investigating whether five-year-old Mandarin-speaking children can understand pre-subject 'only' sentences and

whether prosodic information would help Mandarin speakers (adults and children) to better comprehend this type of sentence.

The current study

In the present study, we aim to answer the question of whether five-year-old Mandarin-speaking children can understand pre-subject 'only' sentences with and without prosodic information when the set of alternatives is explicitly verbally represented in the context, e.g., *The purple panda has the bottle and the cake,* and *The yellow eagle has the cake and the ice cream*, followed by a test sentence *Only the purple eagle has the cake. Is that right?* (For more details, see the Methods section to follow.) Context has been considered an effective tool for referential resolution (Crain and Steedman, 1985. See also Snedeker and Yuan, 2008). Moreover, to extend our knowledge on Mandarin speakers' use of prosodic information during sentence interpretation, three experimental conditions were implemented in the study (i.e., the no accent condition, the head noun accented condition and the modifier accented condition). In the previous chapters, we investigated how Mandarin speakers apply prosodic information to identify the focus in sentences, and it was shown that instead of prosodic information, syntactic information plays an important role in understanding focus. In these studies, we mainly observed how Mandarin speakers apply prosodic information on the sentence level, whereas in the current study, whether Mandarin speakers use prosodic information on the phrase level will be explored. Consider Examples 3 and 4:

3. **Modifier accented condition**:

   Zhiyou ZISE$_F$ de laoying you dangao, shi ma?

   Only PURPLE$_F$ de eagle have cake, Aux Q?

(Only the PURPLE_F eagle has the cake. Is that right?)

4. **Head noun accented condition**:

    Zhiyou zise de LAOYING_F you dangao, shi ma?

    Only purple de EAGLE_F have cake, Aux Q?

    (Only the purple EAGLE_F has the cake. Is that right?)

In Example 3, the modifier *purple* is accentuated in the focus particle phrase (This condition is hereafter referred to as the **modifier accented condition.**), whereas in Example 4, the accent is on the head noun *eagle* of the focus phrase. (This condition is hereafter referred to as the **head noun accented condition.**) Regarding these two examples, the focus is marked by the focus particle 'only' and additional prosodic information (i.e., pitch, duration and intensity). The purpose of placing the accent either on the modifier or on the head noun of the 'only' phrase is to investigate whether Mandarin speakers could apply prosodic information to identify the contrastive sets in the current discourse. Studying these two conditions in the present study will provide us with more insight regarding whether prosodic information could help Mandarin speakers (adults and children) to identify the focus in a sentence in addition to the information provided by the focus particle 'only'. Following Zhou and his colleagues (2012), we expect that prosodic information would help Mandarin-speaking adults to resolve the ambiguity of pre-subject 'only' sentences, but it may not help Mandarin-speaking children. Moreover, the current study would help us to understand whether five-year-old Mandarin-speaking children can understand pre-subject 'only' sentences, which has not been addressed in previous research.

4.2 Method

4.2.1 Participants

Fifty-four five-year-old Mandarin-speaking children (mean age: 5.05; age range: 5.0–5.11) who were living in Kaohsiung, Taiwan, participated in the study. Their parents reported them as being typically developing children with no known visual or hearing impairments at the time of testing. Nine were excluded from further data analysis: one was an Indonesian–Mandarin bilingual child, and the other eight children only provided *yes* responses, which suggested that they did not understand the task. Thirty-nine Mandarin-speaking adults (mean age: 23; age range: 19–35) who were living in Taipei, Taiwan[22], served as the control group. Each participant was randomly assigned to one of the three experimental conditions: 16 children and 13 adults were tested in the no accent condition, 12 children and 13 adults in the head noun accented condition and 16 children and 13 adults in the modifier accented condition. All participants submitted their informed consent documents prior to testing[23].

4.2.2 Materials and design

Sentence–picture verification task. The experiment consisted of 35 trials for each condition, i.e., the no accent condition, the head noun accented condition and the modifier accented condition. Each condition comprised an introduction trial, two practice trials,

---

[22] All adult participants reported that they had learned English as a foreign language starting at age twelve and had infrequent contact with English at the time of testing.

[23] Parents submitted informed consent documents stipulating that their children could participate in the experiment, and the verbal consent of the child participants was obtained prior to testing.

eight with-'only' test trials, eight with-'only' control trials, eight without-'only' test trials and eight without-'only' control trials. In the with-'only' and without-'only' test trials, the test sentences were false descriptions of the pictures, whereas in the with-'only' and without-'only' control trials, the test sentences were true descriptions of the pictures, which allowed us to balance the number of true and false statements. Following Paterson and his colleagues (2003), the purpose of the without-'only' test trials was to assess whether the participants would interpret the sentences with the focus particle 'only' in the same way as they did the ones without. That is, comparing the responses in the with-'only' test trials with the ones in the without-'only' test trials would show whether the participants ignored the information of the focus particle '*zhiyou'* ('only'). (See also the previous Introduction section.).

To investigate whether five-year-old Mandarin-speaking children can understand pre-subject 'only' sentences with prosodic information, three experimental conditions were included in the current study, i.e., the no accent condition, in which no additional prosodic information was added to the head noun or the modifier (later called the default condition); the head noun accented condition and the modifier accented condition. In addition, in each condition, two different types of sentences (i.e., sentences with the focus particle 'only' and sentences without the focus particle 'only') were included to study whether the adult and child participants ignored the information of the focus particle 'only'. We expected that both the adults and the children would treat the with-'only' test trials and without-'only' test trials differently (Paterson et al., 2003) and provide different answers for them. In Table 4.1, example sentences with the focus particle 'only' are shown for each condition. Table 4.2 presents example sentences without the focus particle 'only' for each condition. The default condition was applied to examine whether Mandarin-speaking children can understand pre-subject 'only' sentences with the focus particle

'*zhiyou*' when no additional prosodic information was given. (See Example 1 as a test sentence in the with-'only' test trials and Example 4 as a test sentence in the with-'only' control trials in Table 4.1.) The head noun accented condition (See Example 2 as a test sentence in the with-'only' test trials and Example 5 as a test sentence in the with-'only' control trials.) and the modifier accented condition (See Example 3 as a test sentence in the with-'only' test trials and Example 6 as a test sentence in the with-'only' control trials.) were used to study whether Mandarin speakers could utilise prosodic information to identify the focus constituents that were associated with the focus particle '*zhiyou*'. In Table 4.2, there are example sentences for the without-'only' test (Example 7 for the default condition, Example 8 for the head noun accented condition and Example 9 for the modifier accented condition) and control trials (Example 10 for the default condition, Example 11 for the head noun accented condition and Example 12 for the modifier accented condition) for each condition.

Table 4.*1* Examples of the **with-'only'** test sentences in the test and control trials for the three conditions

Test trials (eight test sentences for each condition in the experiment)
Example 1: Default condition   Zhiyou zise de laoying you dangao, shi ma?   Only purple ASSOC[24] eagle have cake, Aux Q?   (Only the purple eagle has the cake. Is that right?)   Predicted answer: *No, all of them have the cake.*

---

[24] ASSOC stands for associative, and the function of 'de' is to connect the words 'purple' and 'eagle' (Li and Thompson, 1981, p.113).

Example 2: Head noun accented condition

Zhiyou zise de LAOYING$_F$ you dangao, shi ma?

Only purple ASSOC EAGLE$_F$ have cake, Aux Q?

(Only the purple EAGLE$_F$ has the cake. Is that right?)

Predicted answer (following the prosodic information):

No, the purple panda also has the cake.

Example 3: Modifier accented condition

Zhiyou ZISE$_F$ de laoying you dangao, shi ma?

Only PURPLE$_F$ ASSOC eagle have cake, Aux Q?

(Only the PURPLE$_F$ eagle has the cake. Is that right?)

Predicted answer (following the prosodic information):

No, the yellow eagle also has the cake.

Control trials (Eight test sentences for each condition in the experiment)

Example 4: Default condition

Zhiyou zise de laoying you dangao, shi ma?

Only purple ASSOC eagle have cake, Aux Q?

(Only the purple eagle has the cake. Is that right?)

Predicted answer:

Yes

Example 5: Head noun accented condition

Zhiyou zise de LAOYING$_F$ you dangao, shi ma?

Only purple ASSOC EAGLE$_F$ have cake, Aux Q?

(Only the purple EAGLE$_F$ has the cake. Is that right?)

Predicted answer:

Yes

Example 6: Modifier accented condition

Zhiyou ZISE_F de laoying you dangao, shi ma?

Only PURPLE_F ASSOC eagle have cake, Aux Q?

(Only the PURPLE_F eagle has the cake. Is that right?)

Predicted answer:

Yes

Table 4.*2* Examples of the **without-'only'** test sentences in the test and control trials for the three conditions

Test trials (Eight test sentences for each condition in the experiment)
Example 7: Default condition   Zise de laoying you dangao, shi ma?   Purple ASSOC eagle have cake, Aux Q?   (The purple eagle has the cake. Is that right?)   Predicted answer: *No, the purple eagle has the candy and the cookie.*
Example 8: Head noun accented condition   Zise de LAOYING_F you dangao, shi ma?   Purple ASSOC EAGLE_F have cake, Aux Q?   (The purple EAGLE_F has the cake. Is that right?)   Predicted answer (following the prosodic information):   *No, the yellow eagle has the cake.*
Example 9: Modifier accented condition   ZISE_F de laoying you dangao, shi ma?   PURPLE_F ASSOC eagle have cake, Aux Q?

(The PURPLE$_F$ eagle has the cake. Is that right?)

Predicted answer (following the prosodic information):

No, the purple panda has the cake.

Control trials (Eight test sentences for each condition in the experiment)

Example 10: Default condition

Zise de laoying you dangao, shi ma?

Purple ASSOC eagle have cake, Aux Q?

(The purple eagle has the cake. Is that right?)

Predicted answer:

Yes

Example 11: Head noun accented condition

Zise de LAOYING$_F$ you dangao, shi ma?

Purple ASSOC EAGLE$_F$ have cake, Aux Q?

(The purple EAGLE$_F$ has the cake. Is that right?)

Predicted answer:

Yes

Example 12: Modifier accented condition

ZISE$_F$ de laoying you dangao, shi ma?

PURPLE$_F$ ASSOC eagle have cake, Aux Q?

(The PURPLE$_F$ eagle has the cake. Is that right?)

Predicted answer:

Yes

Now, let us turn to the details of the visual stimuli. Each picture displayed three animals with two objects close to each animal. In the pictures for the with-'only' test trials, one of

the objects was the same for all animals (e.g., cake), and the other object was the same for two of the animals: the one in the middle and the animal on the right or the left. (e.g., ice cream. See Figure 4.1.) The unshared objects with the left and right animals were addressed in the context text, and the one shared by all three animals was mentioned in the target sentence. Thus, the context text for Figure 4.1 was *The purple panda has the bottle and the cake, and the yellow eagle has the cake and the ice cream,* and the target sentence was *Only the purple eagle has the cake. Is that right?* The purpose was to inform the participants of the objects' labels and to ensure that the participants could identify the objects. In the pictures for the with-'only' control trials, the object pairs that the animal in the middle had were different from the object pairs that the animal on the left and right hand, but one of the objects that the middle animal had was the same as the one that the animal on the right or left had. (See Figure 4.2.). Again, the same idea that applied to the test trials applied to the control trials: the labels of the objects in the picture were mentioned in the context text. As such, the participants could identify the objects' labels in the picture. Further, eight extra pictures that portrayed the animal in the middle having only one object were created as additional with-'only' control trials. (See Figure 4.3.) These pictures were intended as supplementary for the with-'only' control trials, and they were only used for the participants who interpreted the pre-subject 'only' sentences in the same way as they did the pre-object 'only' sentences. For example, instead of responding *Yes, only the purple eagle has the cake*, they would correct objects, such as *No, the purple eagle also has a chocolate,* when they were asked to verify the test sentence *Only the purple eagle has the cake. Is that right?*, as shown in Figure 4.2. Therefore, once the experimenter noticed that the participants interpreted the pre-subject 'only' sentences in the same way as they did with the pre-object 'only' sentences, instead of presenting the stimuli as in Figure 4.2, we showed the pictures as in Figure 4.3 for the rest

of the control trials in the experiment. By doing so, the number of true and false responses were balanced for these types of participants who interpreted the pre-subject only sentences as the pre-object only sentences.

Figure 4.1 An example of the visual stimuli for the with-'only' test trial, in which the animal in the middle has two objects

Figure 4.2 An example of the visual stimuli for the with-'only' control trial

Figure 4.3 An example of the visual stimuli for the additional with-'only' control trial, in which the animal in the middle has only one object

Examples for the visual stimuli of the without-'only' test and control trials are presented below in Figures 4.4 and 4.5. In the pictures for the without-'only' test trials (See Figure 4.4.), the target animal does not have the target object. That is, the purple eagle does not have the cake. (The test sentence for both pictures was *The purple eagle has the cake. Is that right?*) Again, the same idea of how the stimuli were designed for the with-'only' test and control trails (Figure 4.1. and 4.2) was applied here: to ensure that all of the labels of the objects were familiar to the participants, in the pictures of the without-'only' test and control trials, one of the objects that the middle animal had was the same object as one of those that the animal on the right or left had.

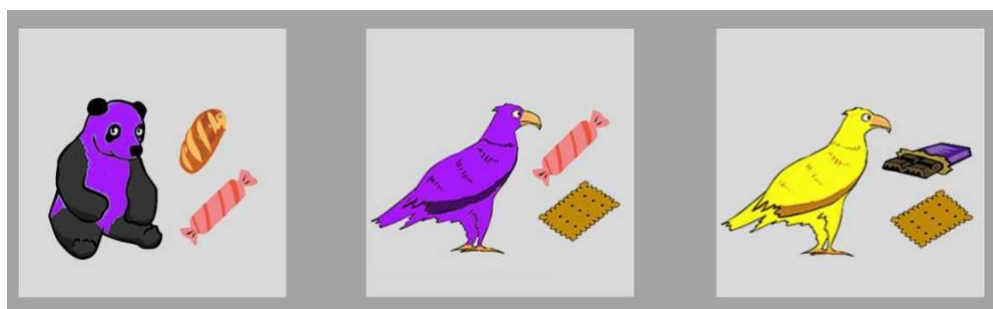Figure 4.4 An example of the visual stimuli for the without-'only' test trial

Figure 4.5 An example of the visual stimuli for the without-'only' control trial

A native female Mandarin speaker recorded all audio stimuli, which were then combined with the visual stimuli into video clips. Each audio stimulus began with a 1s pause, followed by a context text, and before a test sentence was presented, there was a 1s pause between the context text and the test sentence. The length of each audio file was controlled by using sentences with the same number of syllables for each type of stimulus sentence. Each context text lasted 7s. Each test sentence of the with-'only' test trials for the default condition had an average duration of 2.5s, each test sentence of the with-'only' test trials for the head noun accented condition had an average duration of 3.3s and each test sentence of the with-'only' test trial for the modifier accented condition had an average duration of 3.5s. Moreover, each test sentence of the without-'only' test trial for the default condition had an average duration of 2.25s, each test sentence of the without-'only' test trial for the head noun accented condition had an average duration of 3s and each test sentence of the without-'only' test trial for the modifier accented condition had an average duration of 2.95s[25]. No additional prosodic information was added to the subjects of the sentences in the default condition; on the other hand, the duration and amplitude of the head noun of the subjects in the head noun accented condition sentences and the modifier of the subjects in the modifier accented condition sentences were increased.

4.2.3 Procedure

All participants were tested individually either at the National Normal University in Taipei or at the Zhongshan and Zhenbei Public Kindergarten in Kaohsiung. Each participant was

---

[25] The length of the test sentences was controlled: there were 12 syllables in the with-'only' test sentences and 10 syllables in the without-'only' test sentences.

seated in front of a laptop computer, on which the video clips were presented. An experimenter sat beside the participant and presented the video clips. In the introduction trial, 'Anna' was the storyteller and introduced the four animals, i.e., a purple eagle, a purple panda, a yellow eagle and a yellow panda, to the participants. The participants were told that these four animals went to school at the Happy Kindergarten and that their teacher told them that if they came to school regularly, they would get rewards. Meanwhile, Anna told the participants that she would show them what happened in kindergarten in the following test section, and the participants were asked to respond *yes* when the sentence matched the picture or *no* when the sentence did not match the picture and to give a reason if the description of the picture was incorrect. To familiarise the participants with the sentence–picture verification task, the experiment started with two practice trials that required one *yes* and one *no* response. Each experimental trial started with a description of the picture (the context text), followed by a test sentence. (See Figure 4.6.) Once the participants responded, the experimenter started with the next trial. Each test section lasted for approximately 15–20 minutes, depending on the response time. Each response from the participants was audio-recorded, manually transcribed and coded.

Figure 4.6 An illustration of an experimental trial
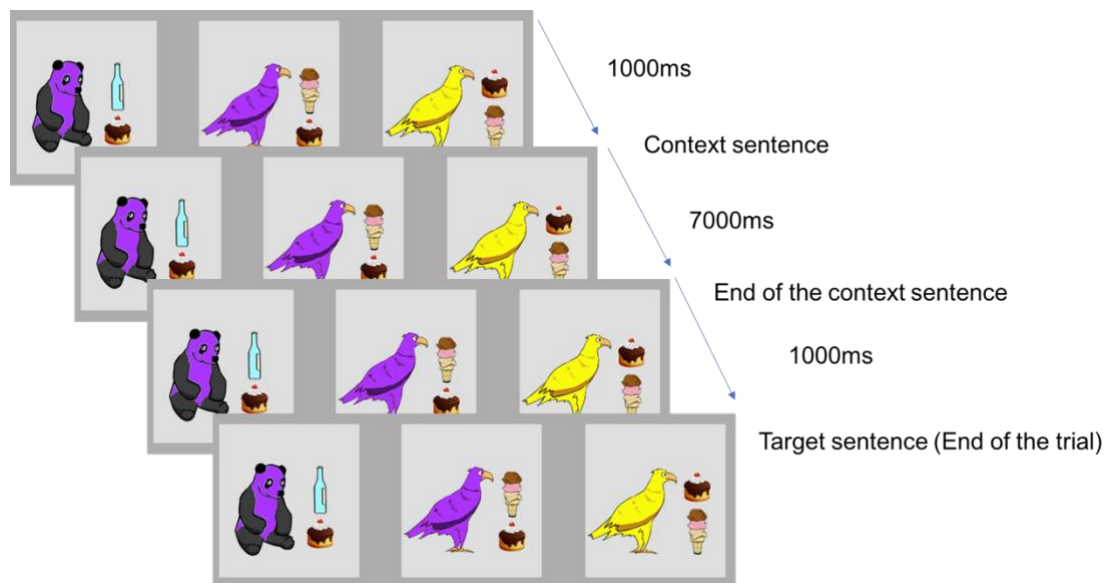
4.3 Results

4.3.1 Scoring of the responses for the with-'only' test trials

Two scoring systems were created, and only the responses for the test trials were included in the scoring system and used for further analysis. One was used to show whether the participant had corrected the subject or the object. (Since in this coding system, the coding criteria solely took into account the participant's consideration of the focus particle 'only', it is referred to as the **only scoring system** hereafter.) The second scoring system was applied to consider whether the participant's correction applied to either the modifier or the head noun. (This coding system is hereafter referred to as the **prosody scoring system** since it revealed whether the participants considered the prosodic information.) The data in the only scoring system were used to investigate whether five-year-old Mandarin-speaking children would associate the focus particle

with the subject of the sentence or whether they would show a general preference for associating the focus particle with the object of the sentence. If the participants corrected the subject, the response was coded as valid and correct, whereas if the participants corrected the object, the response was coded as valid but incorrect. If a response such as *no* was submitted without a specific reason, such as *No, all of them have the cake* or *No, the purple eagle also has the ice cream*, the response was coded as invalid. In total, there were 14 invalid responses among the adult (7 cases) and child (7 cases) participants, and these were excluded from further analysis. The data in the prosody scoring system were used to study whether Mandarin-speaking children and adults could apply prosodic information to identify the focus element. The responses for the default condition were only taken into account in the only scoring system. The responses for the modifier accented and head noun accented conditions were first coded in the only scoring system. To exemplify, take Figure 4.1 with the test sentence *Only the purple eagle has the cake. Is that right?*. First, corrections of the subject in the target test sentences were coded as valid and correct (e.g., *No, all animals have the cake.*), while the corrections of the object were coded as valid but incorrect (e.g., *No, the purple eagle also has an ice cream.*) in the only scoring system.

Figure 4.1 (repeated)

Second, if the participants in the modifier accented or head noun accented condition gave only a correction for the alternative subject (e.g., *No, the purple panda also has the cake* or *No, the yellow eagle also has the cake*.) and for both subjects (e.g., *No, the purple panda and the yellow eagle also have the cake*.), these responses were coded as correct and valid in the only scoring system, as well.

Next, the responses that were coded as correct and valid in the only scoring system were further considered in the prosody scoring system. In the prosody scoring system, the corrections were coded as valid and correct if the prosodic information was considered in the sentence's interpretation. Concerning our example, the correction of the head noun in the head noun accented condition was coded as valid and correct (e.g., Test sentence: *Only the purple EAGLE_F has the cake. Is that right*? Response: *No, the purple panda also has the cake*.), but the correction of the head noun in the modifier accented condition was coded as valid but incorrect (e.g., Test sentence: *Only the PURPLE_F eagle has the cake. Is that right?* Response: *No, the purple panda also has the cake.*). Accordingly, the modifier correction in the head noun accented condition was coded as valid but incorrect (e.g., Test sentence: *Only the purple EAGLE_F has the cake. Is that right?* Response: *No, the yellow eagle also has the cake.*), but the correction of the modifier was considered valid and correct in the modifier accented condition (e.g., Test sentence: *Only the PURPLE_F eagle has the cake. Is that right?* Response: *No, the yellow eagle also has the cake.*). Corrections of the modifier and the head noun (e.g., *No, all animals have a cake*.) were scored as invalid in the prosody scoring system.

Overall, there were only 12 valid corrections among all of the participants (adults and children) in the prosody scoring system, as the majority of the corrections included both alternatives: there was one valid but incorrect case from an adult participant (one modifier correction in the head noun accented condition) and 11 valid cases from the

child participants (two incorrect modifier corrections in the head noun accented condition, four correct modifier corrections in the modifier accented condition and five correct head noun corrections in the head noun accented condition). Thus, the overall number of valid corrections in the prosody scoring system was low. (See Table 4.3.) Therefore, the results for the prosody scoring system were not considered for inferential statistics.

Table 4.*3* Corrections coded in the prosody scoring system (among children and adults) in percentages

Experimental conditions	Valid corrections		Invalid corrections		Total
	Modifier correction	Head noun correction	Corrections of both subjects	Corrections of objects	
Modifier accented Condition	1.7%	0%	90.5%	7.8%	100%
Head noun accented Condition	1.5%	2.5%	87%	9%	100%

4.3.2 Results from the only scoring system

Let us first consider the responses in the without-'only' trials. Again, the purpose of including the without-'only' trials in this study was to check whether the participants submitted different responses for the with-'only' trials and the without-'only' trials and, thus, provide evidence that they did not ignore the meaning of the focus particle 'only'

(Paterson et al., 2003). Overall, all of the participants (adults and children) were completely correct in the without-'only' trials across all conditions. Moreover, regardless of the condition, in the without-'only' test trials, all participants corrected what the target subject had, e.g., *No, the (purple) eagle has the candy and the cookie,* while in the without-'only' control trials, they all responded *yes.* Further, our data showed that the participants submitted different responses for the with-'only' test trials and the without-'only' control trails. Regarding the stimuli for the without-'only' control trials and with-'only' test trials, all three animals had the same object (e.g., ice cream). While in the with-'only' test trials, the participants corrected the statement by saying, *No, all three animals have the ice cream*, they responded *yes* in the without-'only' control trials. The adults and children's responses for the without-'only' control trails were all correct, and the same responses were given regardless of the condition. That is, whether the participants were in the default, head noun accented or modifier accented condition, they responded *yes.* This indicated that the participants did address the information of the focus particle 'only' for the with-'only' sentences. Since all participants performed optimally in the without-'only' trials across all three conditions, there was no need for further inferential statistics.

The percentages of subject corrections according to the only scoring system for both adults and children in the three with-'only' test conditions are presented in Figure 4.7. For inferential statistics, a linear mixed-effect model (Baayen et al., 2008) in R (R Core Team, 2016) was used in the R-studio environment (version 1.1.442). The aim was to assess the impact of sentence conditions on age groups and their interactions regarding the number of correct responses in the with-'only' test trials. In addition to these fixed factors, two random effects were included in the model, namely by-subject variance, which was controlled by age variance, and by-item variance for the responses (as the

random slope)[26]. Moreover, following Zhou and his colleagues' findings (2012), the data in the modifier accented condition were used as the baseline for comparison with the responses for the other two conditions, i.e., the default condition and the head noun accented condition. First, we tested whether the number of subject corrections in the modifier accented condition differed significantly between the age groups. Secondly, the children's responses were tested according to whether the number of subject corrections for the modifier accented condition was significantly different from the number of subject corrections for the other two conditions, i.e., the head noun accented and default conditions. Finally, the interaction between the age groups and conditions was analysed to discover whether there were differences between the modifier accented condition and the default condition, as well as between the modifier accented condition and the head noun accented condition, for the two age groups. Additionally, we explored whether the number of subject corrections of each age group for each condition was significantly different from the chance level by using a one-sample Wilcoxon test.

Overall, the child participants' corrections did not significantly differ from those of the adult participants in the modifier accented condition (Group$_{adults}$: $b = 25.107$, $SE = 29.327$, $Z = 0.856$, $p = 0.392$). Moreover, the child participants did not show significant differences in their performance across the different conditions (Condition$_{modifier\ vs.\ default}$: $b = -0.603$, $SE = 4.631$, $Z = -0.120$, $p = 0.896$; Condition$_{modifier\ vs.\ head\ noun}$: $b = -0.749$, $SE = 4.776$, $Z = -0.157$, $p = 0.875$). Further, there were no interaction effects between the conditions and the age groups (Age group$_{children\ vs.\ adults}$ * Condition$_{modifier\ vs.\ default}$: $b = -25.416$, $SE = 29.482$, $Z = -0.862$, $p = 0.389$; Age group$_{children\ vs.\ adults}$ * Condition$_{modifier\ vs.\ head\ noun}$: $b = -11.412$, $SE = 38.092$, $Z = -0.300$, $p = 0.764$). On the other hand, because the adults' performance in the default condition was not optimal, as in the other two conditions, we

---

[26] Response ~ group * condition + (1|age:id) + (1|item)

applied the same model again, but this time, the adult group was set as the baseline; the results showed that there were no differences among the conditions for the adult age group, either.

In addition, the adult participants' performance was optimal for the head noun accented and modifier accented conditions. Performance was above the chance level for both groups in all conditions. (Child participants in the default condition: $p = 0.003$; child participants in the head noun accented condition: $p = 0.017$; child participants in the modifier accented condition: $p = 0.000$. Adult participants in the default condition: $p = 0.014$; adult participants in the head noun accented condition: $p = 0.000$; adult participants in the modifier accented condition: $p = 0.000$.) Overall, five-year-old Mandarin-speaking children did not behave differently from Mandarin-speaking adults, and both Mandarin-speaking children and adults performed equally well in all three conditions.

Figure 4.7 The overall subject correction percentages of Mandarin-speaking adults and children for the with-'only' test sentences of the three conditions
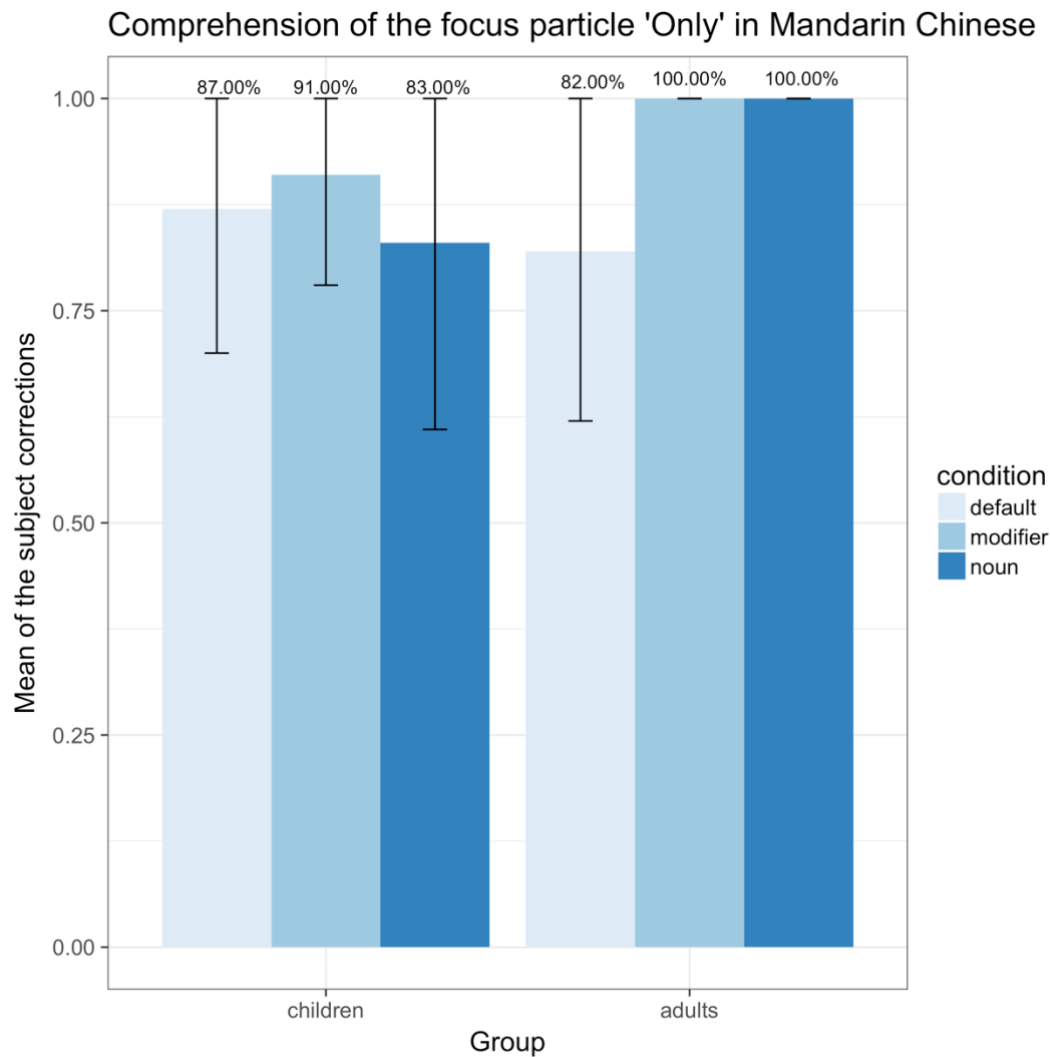
4.4 Discussion

The sentence–picture verification task was applied in the current study to investigate whether five-year-old Mandarin-speaking children could understand pre-subject 'only' sentences, as well as whether they could apply prosodic information to identify the focus and construct the corresponding alternatives of the pre-subject 'only' sentences. The results showed that the children preformed above the chance level in all experimental

conditions and that the children's performance was not significantly different from adult-like performance in all conditions. The present results are in contrast to what has been found in previous cross-linguistic research that has shown that children up to the age of six years struggle to comprehend pre-subject 'only' sentences (English: Crain et al., 1994; Paterson et al., 2003. Portuguese: Costa and Szendrői, 2006. Mandarin: Yang, 2002).

The current findings are somewhat incompatible with the assumptions of Paterson et al. (2003) and Müller et al., Schulz and Höhle (2011) that state that children cannot establish the set of alternatives that are relevant for the interpretation of the sentences with the focus particle 'only'. Müller and her colleagues (2011) replicated Paterson and his colleagues' (2003) study with six-year-old German-speaking children. In their first study, Müller and her colleagues (2011) examined children's comprehension of the focus particle 'only' using the forced-choice picture-selection task. According to their results, German-speaking children failed to establish the alternatives in the pre-subject and pre-object 'only' sentences; hence, they could not understand the sentences with the focus particle 'only' correctly, as shown by the fact that the child participants selected the same picture for the sentences with the focus particle 'only' and for those without this particle. Their results were the same as what Paterson and his colleagues (2003) found. Moreover, based on their first results, Müller and her colleagues further investigated whether children had difficulties in addressing information that was not verbally introduced in the discourse by using a felicity judgement task in their second experiment. Four different types of the visual stimuli that were used in the first study were applied in the second study, together with the test sentences without the focus particle 'only', e.g., *The fireman is holding a hose*. These four pictures varied in their degrees of visual information complexity. The participants were asked to verify whether the target sentences matched the visual stimulus. The authors assumed that the picture with the least amount of visual

information would be the most acceptable one (E.g., in the picture, there were two firemen, and one of them was holding a hose.), and the picture with the most complex visual information would be the least acceptable one (E.g., in the picture, there were two firemen, and one of them was holding a hose and a ladder, and the other one was holding a hose, too.) when the sentence mentioned only part of the information provided in the picture. The results confirmed the predictions that the visual complexity information did affect the participants' sentence judgements since the sentences with the most complex visual information were the least acceptable, whereas those with the least amount of visual information were the most acceptable. This especially held for the child participants: they expected that the given scene had to maximally match with the informed scenario. That is, if the set of alternatives for the sentence with the focus particle in the given scene was verbally provided to the children, they were better able to understand the sentence. Again, this finding helped to indicate that due to the infelicitous information in the first experiment, German-speaking children could not reach adult-like performance for the pre-subject and pre-object 'only' sentences.

In contrast, in the present experiment—in which the sets of alternatives were explicitly verbally provided via the context text—five-year-old Mandarin-speaking children could represent the set of alternatives when presented with pre-subject 'only' sentences: they provided eight responses for the with-'only' test trails that were different from the eight responses for the without-'only' control trials. In the condition where the sentence included the focus particle 'only', e.g., *Only the purple eagle has the cake. Is that right?*, Mandarin speakers (adults and children) recognised the correct set of alternatives i.e., that the yellow eagle and the purple panda also had the cake, and corrected the statement as *No, all animals have the cake.* On the other hand, in the condition where the target sentence did not include the focus particle 'only', e.g., *The purple eagle has the cake.*

Is that right?, presented with a picture where all three subjects had the cake and something else, the participants responded *yes.* Interestingly, some participants, especially the child participants, were in favour of responding with the focus particle 'only' when they were asked to verify the sentences without this focus particle. The participants integrated the focus particle 'only' into their responses to emphasise what the animal had, e.g., *No, the purple eagle only has the candy and the cookies.* This again provides evidence supporting the fact that children know the lexical meaning and the syntactic scope of the focus particle 'only': they corrected the subject in the pre-subject 'only' sentences and added the focus particle 'only' to stress the object of the sentence. This provides evidence against the syntactic account that assumes that children are not sensitive to the scope of the focus particle 'only' (Crain et al., 1994; Notley et al., 2009; Zhou et al., 2012).

On the other hand, our results did not coincide with those of Yang's (2002) study. In her study, only Mandarin-speaking children equal to or above the age of eight years could comprehend the pre-subject 'only' sentences, whereas in the current study, five-year-old children could already understand the pre-subject 'only' sentences. Yang suggested that the mismatch in information statuses (topic and focus) in the pre-subject 'only' sentences resulted in the speakers struggling to correctly interpret this type of focus structure (the discourse function account, also see Müller et al., 2013). Similar to the studies by Peterson et al. and Müller et al., in Yang's study, the context was not verbally explicitly provided: in each test trial, three different scenario pictures were presented, one of which depicted only a boy carrying a bucket and no one else carrying a bucket, another depicted the boy carrying a bucket and a school bag and the last depicted the boy carrying a bucket and a school bag while playing soccer. The participants were asked to verify these pictures with four different types of sentences, i.e., the pre-subject sentence with the focus particle '*jiu*', the pre-subject sentence with the focus particle '*zhi*', the pre-

object sentence with the focus particle '*jiu*' and the pre-object sentence with the focus particle '*zhi*'. On the one hand, the context was not verbally explicitly provided, and this might have influenced the performance of the younger participants (the four- to six-year-olds). On the other hand, since the two pre-subject 'only' test sentences and two pre-object 'only' test sentences were provided to the same participants, this may have made it more difficult for the child participants to verify the four different types of sentences within one study. In the current study, no pre-object 'only' sentences were used in such a way that interference could have arisen between the two sentence types.

It has been argued that adults, but not children, use prosodic information to resolve structural ambiguities (Snedeker and Trueswell, 2001; Choi and Mazuka, 2003; Snedeker and Trueswell, 2003). However, a later online study that Snedeker and Yuan (2008) conducted indicated that English-speaking children could utilise prosodic information to resolve ambiguity in a sentence such as *You can feel the frog with the feather,* but with a 500ms delayed processing time compared to adults. Similarly, a study in Mandarin by Zhou and his colleagues (2012) examined whether four- to five-year-old Mandarin-speaking children could apply prosodic information to resolve ambiguity in pre-subject 'only' sentences. Their results indicated that although children showed adult-like performance in their looking patterns, they needed more time than adults for the integration of syntax and prosody (400ms delay). For example, in their behavioural responses, the child participants constantly corrected the modifier, irrespective of the prosodic information. However, in the eye gaze data, the children showed adult-like performance, as they looked more to the head noun when it was accented, but with a delay of 400ms compared to the adults. Moreover, the authors of this study argued that Mandarin-speaking children have a default interpretation for the pre-subject 'only' sentence *Only Xiaoming's clock is yellow,* in which they associated the focus particle 'only'

with the modifier. Our results are partially in line with what Zhou and his colleagues (2012) found. In the current experimental design, if the speakers followed the prosodic information, they would have corrected the modifier in the modifier accented condition or the head noun in the head noun accented condition. However, in the current findings, in total, there were only 12 participants (one in the adult data and 11 in the child data) in the prosody condition who corrected either the modifier or the head noun. For the one adult response, the modifier was corrected in the head noun accented condition, and for the 11 child responses, the modifier was corrected twice in the head noun accented condition and four times in the modifier accented condition, and the head noun was corrected five times in the head noun accented condition. The number of corrections following the prosodic information was too low to draw any further conclusions, but there was no indication of a preferred association between the focus particle and the modifier, as Zhou et al. (2012) reported. Whether this is related to the different types of modifiers used in the two studies (NPs in Zhou et al. and adjectives in the current study) is a question for further research.

Further concerning the prosodic information, Mandarin-speaking adults performed better in the prosody conditions, i.e., they provided more subject corrections in the modifier accented and head noun accented conditions than in the default condition. The additional prosodic information on the modifier or the head noun of the subject might have helped to overcome the default interpretation of the focus on the objects (Szendrői et al., 2017; Chen et al., 2019). However, the differences among the conditions were not proven by the statistical analysis; thus, more research regarding how Mandarin speakers represent prosodic information and utilise it for focus assignment is necessary.

In summary, five-year-old Mandarin-speaking children can understand pre-subject 'only' sentences when an explicit verbalised context is provided. Additionally,

there is a non-significant indication that prosodic information helps adults to associate the focus particle with the subject. To shed more light on the issue raised in the current study, several improvements could be made in the future. In the current study, only an explicit task was included. Following the study in the previous chapter, an implicit method—such as an eye-tracking experiment—might provide better insight into how Mandarin speakers (adults and children) comprehend pre-subject 'only' sentences, as well as how they exploit prosodic information to identify the focus element (Zhou et al., 2012; also see Höhle et al., Fritzsche and Müller, 2016). Further, in our study, only five-year-olds were tested, and to further scrutinise the developmental trajectory of the focus particle in children, children as young as four years old could be examined. (In Yang's study (2002), four-year-old Mandarin speakers treated the pre-subject 'only' sentences as the pre-object 'only' sentences.)

5 General Discussion

The studies presented in this thesis aimed at disentangling the issue of how tone language speakers and intonation speakers (children and adults) acquired various linguistics cues, i.e., prosodic information, syntactic information, focus particle- only, to comprehend sentences focus in a cross-linguistics perspective. A number of studies have been done with intonation language speakers; however, less has been carried out with tone language speakers. Hence, within this thesis a series of thoroughly controlled experiments were conducted with intonation language speakers (German-speaking adults and children) as well as with tone language speakers (Mandarin-speaking adults and children). Second, eye-tracking methods has been considered as a tool to directly reveal the ongoing interpretation process. It can be used to demonstrate where the speaker directs the attention to when presented with a sentence and a visual stimulus; therefore, this method will help us to advance our knowledge in how children and adults process focus information.

Four experimental studies were carried out. The first one examined how Mandarin speakers (children and adults) apply prosodic information to identify sentence focus. The second and third study investigated how Mandarin and German speakers (children and adults) make use of prosodic information and syntactic information to recognize sentence focus. The fourth study tested whether five-year-old Mandarin speaking children could understand pre-subject only sentences and whether prosodic information could help them to better understand this kind of sentences. In the following, a summary of the findings and their theoretical implications of each study will be provided, followed by some recommendations for the future research and an overview of the main milestones that were achieved through this thesis.

5.1 Study Summary

5.1.1 Acquisition of Prosodically-Marked Focus

In the thesis, two studies were conducted to examine how tone language- Mandarin-speakers and intonation language- German- speakers comprehend prosodically-marked focus. In **Chapter Two**, a task originally designed by Szendrői and her colleagues (2017) - a sentence-picture verification task - was carried out to examine whether three- to five-year-old Mandarin-speaking children as well as Mandarin-speaking adults could apply prosodic information to identify the focus in sentences. Each participant would either hear a subject-accented sentence (Subject-accented condition) or an object-accented sentence (Object-accented condition) and they were asked to verify the information on pictures they saw together with the sentences they heard. Contrary to Chen's (1998) findings, our results provided evidence showing that children from early onwards (three-year-olds) could make use of the specific language cues in their ambient language, that is, word-order in Mandarin instead of prosodic information to identify focus (Szendrői et al., 2017). That is, instead of following the prosodic information to correct either the sentence subject or the sentence object, participants (children and adults) mostly corrected the sentence object which is located in the default focus position in Mandarin (Xu, 2004). Much more studies have been paying attention to how Mandarin-speaking children and adults could apply the prosodic information to produce focus (Chen and Braun, 2016; Chen and Gussenhoven, 2008; Ouyang and Kaiser, 2014; Xu, 1999) and less has been done with the tone language speakers of how they comprehend the prosodically-marked focus. Therefore, in the current study, we aimed at bringing better understanding of whether Mandarin speakers (both children and adults) were sensitive to the prosodic information

to identify the sentence focus. In contrast to the previous findings that Mandarin-speaking children were more sensitive to prosodic information than Mandarin-speaking adults, our results suggested that both Mandarin-speaking children and adults were not sensitive to the prosodic information and children replied more on the word-order information than on the prosodic information to identify the sentence focus. These findings add to a growing body of literature on how tone language learners as well as Topic-prominent language learners comprehend prosodically-marked focus and the article is published in the Journal of Psycholinguistics Research (2019).

To add more understanding with the intonation language learners, in **Chapter 3**, not only Mandarin speakers (tone language) but also German speakers (intonation language) were studied to further test the assumption that children have adult-like performance in understanding sentence focus by identifying language specific cues in their mother tongue from early onwards. Second, regarding focus being the indication of the alternative (Rooth, 1992), the visual word paradigm (VWP) was considered as a more sensitive experimental method for the research question carrying out here. On the one hand, the VWP has been applied to observe how the parser computes a temporally ambiguous structure (Tanenhaus et al., 1995; Snedeker and Trueswell, 2003; Trueswell et al., 1999; Snedeker and Trueswell, 2004). On the other hand, it is agreed that for both adults and children, where their eyes fixate is where their attention is (Sekerina et al., 2004; Brandt-Kobele and Höhle, 2010; Höhle, Fritzshce and Müller, 2016). Therefore, in this study, the same paradigm - the sentence-picture verification task - as in the first study was employed; in addition, eye-tracking data were obtained while participants solved the task. Following the findings in Chapter 2, we hypothesized that Mandarin speaker (both children and adults), instead of adhering to the prosodic cues would follow word-order information to identify the sentence focus. On the other hand, German speakers would be

able to apply the additional prosodic information on sentence subject and sentence object to identify the focus (Szendrői et al., 2017). Again, the results corroborate with the findings in Chap 2 that Mandarin speakers were aware of the specific language cues, i.e., the word-order in an early stage (in this study, five years old). Both Mandarin-speaking adults and Mandarin-speaking children preferred correcting the sentence object over the sentence subject in the prosodically-marked focus conditions, i.e., in the subject-accented condition as well as in the object-accented condition. However, the results with the German speakers were not as clear as with the Mandarin speakers. On the one hand, German-speaking adults performed over chance level both in the subject-accented condition (78% of the time corrected the sentence subject) and in the object-accented condition (89% of the time corrected the sentence object). On the other hand, German-speaking children were better in the subject-accented condition (66% of the time corrected the sentence subject) than in the object-accented condition (40% of the time corrected the sentence object). In German, as in Mandarin and other languages, the default focus position is in the most deeply embedded sentence position (Xu, 2004; Reinhard, 2004). Therefore, for the sentence object in the object-accented condition sentence, it carried not only the nuclear stress following the default focus position, but also it contained the additional prosodic information. One would expect that participants would perform better in the object-accented condition than in the subject-accented condition since there were two hints provided at the same time- the prosodic information and the default focus position - as the sentence focus marker. However, this extra information might confuse the younger participants. In this study, the five-year-old German-speaking children submitted the most double corrections in the object-accented condition among the other conditions. It could be that for the German-speaking children the focus assignment in the object-accented condition is less clear and this finding was comparable

to the six-year-old German-speaking children's results reported in Szendrői and her colleagues' study (2017). Future work would be needed to further investigate how prosodic information interacts with the default focus position.

5.1.2 Acquisition of Syntactically-Marked Focus

5.1.2.1 Acquisition of the cleft construction

On the one hand, there were numbers of theoretical studies in Mandarin discussing both the grammaticality and the functionality of cleft constructions but few empirical studies on how Mandarin-speaking children acquire this kind of construction (Theoretical studies: Tang, 1983; Cheng, 2008; Paul and Whitman, 2008; Hole, 2011; Empirical studies: Chen 1998). On the other hand, the characteristics of the cleft constructions in German were not well described and no systematic study on the acquisition of the cleft constructions in German children is available (Dufter, 2009). Therefore, in the study described in **Chapter 3**, in addition to the prosodic information, the syntactic information, i.e., cleft constructions, was further included to enhance our knowledge in how tone language speakers and intonation language speakers identify sentence focus.

Mandarin-speaking adults performed above chance level in the object-pseudocleft and the subject-cleft conditions whereas Mandarin-speaking children performed different from chance only in the object-pseudocleft condition but performed at chance level in the subject-cleft condition. The results are consistent with the previous findings in Chen's study (1998) that object-pseudocleft sentences are easier for Mandarin-speaking children than subject-cleft sentences. The reason for this may be that there are two cues for recognizing the focus in the object-pseudocleft condition, i.e., the default

focus position and the focus marker- *SHI* (Xu, 2004; Tang, 1980). More, in addition to the off-line measurements, the eye-tracking method was implemented in the study in which the looking patterns of the participants were recorded. Both Mandarin-speaking adults and Mandarin-speaking children looked more to the subject alternative in the subject-cleft condition than in the object-pseudocleft condition. That is, while processing the subject-cleft sentences, Mandarin speakers (adults and children) were aware of the syntactically-marked focus as shown by their eye looking patterns. However, Mandarin-speaking children preferred the default focus position to the marked structure and in the end they corrected the sentence object instead of the sentence subject. Taken together, eye-tracking method has strengthened our knowledge in how the parser direct their attention for recognizing focus: while Mandarin-speaking children performed worse than Mandarin-speaking adults in the subject-cleft condition, children's eye-looking patterns mirrored the adults' eye-looking patterns in the subject-cleft condition.

Regarding the German results, first, German-speaking adults performed at chance level in the subject-cleft condition as well as in the object-cleft condition. More, their performance was also in line with their eye looking patterns as they did not look more to the subject alternative neither in the subject-cleft condition nor in the object-cleft condition. Further, while the German-speaking children had adult-like performance in the subject-cleft condition (51% of the time corrected the sentence subject), they only corrected the sentence object 26% of the time in the object-cleft condition. It is assumed that German-speaking children applied a default strategy to identify focus and further to verify the target sentences in the object-cleft condition. That is, they concerned the default word-order information while correcting the target in the object-cleft sentences. Therefore, instead of correcting the alternative object, they corrected the alternative subject. Overall, our findings associate with what Dufter (2009) noted in his corpus study:

low occurrence of the cleft construction in German might be due to multiple means of addressing focus in German, such as word-order, adding focus particle, etc. Therefore, understanding the cleft construction might be inconstant in German concerning the adults' data with the clefted sentences. Additionally, we assumed that there were more possibilities of marking focus in German; therefore, it would take German-speaking children longer time to manage both the prosodic and the syntactic information for identifying focus in sentences. This study is the very first empirical study in understanding how German apprehend and acquire the cleft construction and more studies are required to gain a deeper understanding of the cleft construction in German.

5.1.2.2 Acquisition of focus particle- only

In the last study in Chapter 4, an issue of whether five-year-old Mandarin-speaking children could understand the pre-subject only sentence was carried out and again whether prosodic information would help them to better understand this kind of sentences. Cross-linguistically it has been indicated that only children up-to-school age could manage the pre-subject only sentences while three-year-olds could already interpret the pre-object only sentences correctly (Berger and Höhle, 2012; Crain et al., 1994; Müller et al, 2011; Yang, 2002). Our results showed, however, a different picture that with explicit verbalized context five-year-old Mandarin-speaking children could comprehend the pre-subject only sentences. Secondly, concerning our sentence stimuli, the subject NP comprised of an adjective and a noun phrase; hence, the extra information, i.e., the adjective, might raise children's attention toward the focus particle- only and further corrected the sentence subject. Last but not least, concerning, our previous research, we hypothesized that prosodic information would only help Mandarin-speaking

adults but not children to better understand the pre-subject only sentences and this was granted in the results. Overall, this study presented that Mandarin-speaking children (five-year-old) could comprehend the pre-subject only sentences. This is an important finding in the understanding of the importance of how the explicitly verbalized context would enhance younger speakers' understanding of the pre-subject only sentence.

5.2 Future Directions

One of the important issues in children language acquisition is acquisition of information structure. This process involves several complex phrases: children, first, would need to relate the lexical knowledge as well as the phonological information to the semantic and pragmatic characteristics in their ambient language and more, the overall message would be represented in the mental grammar. They would further engage the knowledge to communicate with the other speakers. Meanwhile, their cognitive performance is still strengthening to help to mingle the whole process. Even, after several decades of research, it remained unclear how children pick up this immense practice. In this project, the issue of how children acquire different linguistics cues, i.e., prosodic information and syntactic information, in a cross-linguistic perspective has been examined. Regarding prosodic information, German-speaking children, intonation language learners, were sensitive to the prosodically-marked focus, whereas Mandarin-speaking children, tone language learners, could make use of the syntactic information to identify focus. The findings of this project suggests that different language learners- tone language learners as well as intonation language learners- from early onwards were sensitive to their ambient language features of signaling focus (Szendrői et al, 2017). More, although Mandarin-speaking children did not have the adult-like performance in understanding the

syntactically-marked focus which was carried out by the cleft construction, their eye looking patterns mirrored the adults' eye looking patterns. That is, Mandarin-speaking children were sensitive to the syntactic cue in the moment; however, the default word-order information plays a more important role than the focus marker for the tone language learners as well as Topic-prominent language learners. Further, it seems that the syntactically-marked focus indicated by the cleft construction was not valid in German as the German-speaking adults did not performed over chance level neither in the subject-cleft construction nor in the object-cleft construction (Dufter, 2009). This was the first empirical study that investigated the interpretation of the cleft construction in German and more research is needed. In addition, in this project, we were only able to assess two languages, i.e., Mandarin and German, and more cross-linguistic studies are in need. In our final study, we only included the off-line measure and it will be advantageous if the online measurements, such as the eye-tracking method, could be added in.

5.3 Concluding Remarks

As mentioned in the Introduction, this project was conducted in order to investigate how children and adults exploit prosodically-marked and syntactically-marked focus across different languages and whether eye-tracking method would be beneficial in understanding the issued stated above. In conclusion, we have obtained comprehensive results proving that either tone language learners or intonation language learners acquire their language specific cues for marking focus in an early stage: Mandarin-speaking children were sensitive to the syntactically-marked focus while German-speaking children could make use of the prosodic information to identify focus. More, eye-tracking method has helped us in advancing our knowledge in how Mandarin-speaking children processed the syntactic cues even though they could not produce the adult-like performance. Further, the pre-subject only study underlined the importance of the explicit verbalized context and the results demonstrated that Mandarin-speaking children could understand the pre-subject only sentences at the age of five.

Reference

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.

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.

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.

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(8), 1910-1925.

Chao, Y. R. (1968). A grammar of spoken Chinese.

Chen, A. (2011). Tuning information packaging: Intonational realization of topic and focus in child Dutch. Journal of Child Language, 38(5), 1055-1083.

Chen, H.-C., Szendrői, K., Crain, S., & Höhle, B. (2019). Understanding prosodic focus marking in Mandarin Chinese: data from children and adults. Journal of Psycholinguistic Research, 48(1), 19-32.

Chen, S. H. E. (1998, Dec). 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 2006,

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

Cheng, L. L.-S. (2008). Deconstructing the shi… de construction.

Choi, Y., & Mazuka, R. (2003). Young children's use of prosody in sentence parsing. Journal of Psycholinguistic Research, 32(2), 197-217.

Chomsky, N. (1969). Aspects of the Theory of Syntax (Vol. 11). MIT press.

Corrêa, L. M. S. (1995). An alternative assessment of children's comprehension of relative clauses. Journal of Psycholinguistic Research, 24(3), 183-203.

Costa, J., & Szendrői, K. (2006). Acquisition of focus marking in European Portuguese. The acquisition of syntax in Romance languages. Amsterdam & Philadelphia: John Benjamins, 319-330.

Crain, S., Ni, W., & Conway, L. (1994). Learning, parsing and modularity. Perspectives on sentence processing, 443-467.

Crain, S., Philip, W., Drozd, K., Roeper, T., & Matsuoka, K. (1992). Only in child language. Unpublished manuscript, University of Connecticut, Storrs, CT.

Crain, S., & Steedman, M. (1985). On not being led up the garden path: The use of context by the syntactic processor. Natural language parsing: Psychological, computational and theoretical perspectives, Cambridge University Press, New York.

Cutler, A., & Swinney, D. A. (1987, Feb). Prosody and the development of comprehension. Jounal of Child Language, 14(1), 145-167.

De Villiers, J. G., De Villiers, J., De Villiers, P. A., & DeVilliers, P. A. (1978). Language acquisition. Harvard University Press.

Declerck, R. (1984). The pragmatics of it-clefts and wh-clefts. Lingua, 64(4), 251-289.

Diessel, H., & Tomasello, M. (2005). A new look at the acquisition of relative clauses. Language, 882-906.

Dimroth, C., & Narasimhan, B. (2012). The acquisition of information structure. In The expression of information structure (pp. 319-362). De Gruyter Mouton.

Dufter, A. (2009). Clefting and discourse organization: Comparing Germanic and Romance. Focus and background in Romance languages, 83-121.

Feldhausen, I., & Vanrell, M. (2014). Prosody, focus and word order in Catalan and Spanish. An Optimality Theoretic approach. 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, C. (1993). German intonational Patterns. In German intonational Patterns. De Gruyter.

Féry, C., & Krifka, M. (2008). Information structure. Unity and diversity of languages. Amsterdam: John Benjamins, 123-136.

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

Friedmann, N., & Novogrodsky, R. (2004). The acquisition of relative clause comprehension in Hebrew: A study of SLI and normal development. Journal of Child Language, 31(3), 661-681.

Gussenhoven, C. (1984). On the Grammar and Semantics of Sentence Accents. In On the Grammar and Semantics of Sentence Accents. De Gruyter Mouton.

Gussenhoven, C. (2008). Types of focus in English. In Topic and focus (pp. 83-100). Springer.

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

Hedberg, N. (2000). The referential status of clefts. Language, 891-920.

Hedberg, N. A. (1990). Discourse pragmatics and cleft sentences in English University of Minnesota].

Heizmann, T. (2007). Children's acquisition of exhaustivity in clefts. Proceedings of the 31st annual Boston University Conference on Language Development., Somerville.

Höhle, B., Berger, F., & Sauermann, A. (2016). Information Structure in First Language Acquisition. In The Oxford Handbook of Information Structure (pp. 562-580). Oxford University Press.

Höhle, B., Fritzsche, T., & Müller, A. (2016). Children's comprehension of sentences with focus particles and the role of cognitive control: An eye tracking study with German-learning 4-year-olds. PloS one, 11(3), e0149870.

Hole, D. (2011). The deconstruction of Chinese shì… de clefts revisited. Lingua, 121(11), 1707-1733.

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

Hornby, P. A., & Hass, W. A. (1970). Use of contrastive stress by preschool children. Journal of Speech and Hearing Research, 13(2), 395-399.

Hu, S., Gavarró, A., Vernice, M., & Guasti, M. T. (2016). The acquisition of Chinese relative clauses: contrasting two theoretical approaches. Journal of Child Language, 43(1), 1-21.

Huber, S. (2004). Es-Clefts und det-Clefts: Zur Syntax, Semantik und Informationsstruktur von Spaltsätzen im Deutschen und Schwedischen.

Hupet, M., & Tilmant, B. (1989). How to make young children produce cleft sentences. Journal of Child Language, 16(2), 251-261.

Jackendoff, R. S. (1972). Semantic interpretation in generative grammar.

Jespersen, O. (1949). A modern English grammar on historical principles. Part VII. Munksgaard.

Kiss, K. É. (1998). Identificational focus versus information focus. Language, 74(2), 245-273.

Ladd, D. R. (1996). Intonational phonology. Cambridge University Press.

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

Lambrecht, K. (1996). Information structure and sentence form: Topic, focus, and the mental representations of discourse referents (Vol. 71). Cambridge university press.

Lambrecht, K. (2001). A framework for the analysis of cleft constructions.

Lempert, H., & Kinsbourne, M. (1980). Preschool children's sentence comprehension: Strategies with respect to word order. Journal of Child Language, 7(2), 371-379.

Li, C. N., & Thompson, S. A. (1976). Subject and topic. Academic Press New York.

Li, C. N., & Thompson, S. A. (1989). Mandarin Chinese: A functional reference grammar. Univ of California Press.

MacWhinney, B., & Bates, E. (1978). Sentential devices for conveying givenness and newness: A cross-cultural developmental study. Journal of verbal learning and verbal behavior, 17(5), 539-558.

Müller, A., Höhle, B., & Schulz, P. (2013). The interaction of focus particles and information structure in acquisition. Language acquisition and development: Proceedings of GALA,

Müller, A., Schulz, P., & Höhle, B. (2011). Pragmatic children: How German children interpret sentences with and without the focus particle only. Experimental pragmatics/semantics, 79-100.

Musan, R. (2010). Informationsstruktur. Winter.

Narasimhan, B., & Dimroth, C. (2008). Word order and information status in child language. Cognition, 107(1), 317-329.

Narasimhan, B., & Dimroth, C. (2018). The influence of discourse context on children's ordering of "new" and "old" information. Linguistics Vanguard, 4(s1).

Notley, A., Zhou, P., Crain, S., & Thornton, R. (2009). Children's interpretation of focus expressions in English and Mandarin. Language Acquisition, 16(4), 240-282.

Ouyang, I. C., & Kaiser, E. (2014, Feb 7). 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.

Paterson, K. B., Liversedge, S. P., Rowland, C., & Filik, R. (2003). Children's comprehension of sentences with focus particles. Cognition, 89(3), 263-294.

Paterson, K. B., Liversedge, S. P., White, D., Filik, R., & Jaz, K. (2006). Children's interpretation of ambiguous focus in sentences with" only". Language Acquisition, 13(3), 253-284.

Paul, W., & Whitman, J. (2008). Shi… de focus clefts in Mandarin Chinese.

Prince, E. F. (1978). A comparison of wh-clefts and it-clefts in discourse. Language, 883-906.

R Development Core Team. (2016). R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing.

Reeve, M. (2011). The syntactic structure of English clefts. Lingua, 121(2), 142-171.

Reinhart, T. (1981). Pragmatics and linguistics: An analysis of sentence topics in pragmatics and philosophy I. Philosophica anc Studia Philosophica Gandensia Gent, 27(1), 53-94.

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

Reinhart, T. (2006). Interface strategies. OTS working papers in linguistics.

Repp, S., & Drenhaus, H. (2015). 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.

Rooth, M. E. (1985). Association with Focus (Montague Grammar, Semantics, Only, Even) University of Massachusetts Amherst].

Sekerina, I. A., Stromswold, K., & Hestvik, A. (2004). How do adults and children process referentially ambiguous pronouns? Journal of Child Language, 31(1), 123-152.

Selkirk, E. (1995). Sentence prosody: Intonation, stress, and phrasing. The handbook of phonological theory, 1, 550-569.

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. In C. T. J. Huang, Y. H. A. Li, & A. Simpson (Eds.), The Handbook of Chinese Linguistics (pp. 100-125). John Wiley & Sons.

Snedeker, J., Thorpe, K., & Trueswell, J. (2001). On choosing the parse with the scene: The role of visual context and verb bias in ambiguity resolution. Proceedings of the Annual Meeting of the Cognitive Science Society,

Snedeker, J., & Trueswell, J. (2003). Using prosody to avoid ambiguity: Effects of speaker awareness and referential context. Journal of memory and language, 48(1), 103-130.

Snedeker, J., & Trueswell, J. C. (2004, Nov). The developing constraints on parsing decisions: the role of lexical-biases and referential scenes in child and adult sentence processing. Cogn Psychol, 49(3), 238-299.

Snedeker, J., & Yuan, S. (2008). Effects of prosodic and lexical constraints on parsing in young children (and adults). Journal of memory and language, 58(2), 574-608.

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.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995, Jun 16). Integration of visual and linguistic information in spoken language comprehension. Science, 268(5217), 1632-1634.

Tang, T.-c. (1980). Cleft and pseudo-cleft constructions in Chinese: structure, function and constraint. Journal of Taiwan Normal University, 25.

Tang, T.-c. (1983). Focusing constructions in Chinese: Cleft sentences and pseudo-cleft sentences. Studies in Chinese syntax and semantics universe and scope: presupposition and quantification in Chinese, 127226.

Team, R. D. C. (2016). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna. .

Teng, S.-h. (1979). Remarks on Cleft Sentences in Chinese. Journal of Chinese Linguistics, 7(1), 101-114.

Tieu, L., & Kriz, M. (2017). Connecting the exhaustivity of clefts and the homogeneity of plural definite descriptions in acquisition. Proceedings of the 41st Annual Boston University Conference on Language Development,

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(2), 89-134.

Vallduví, E., & Engdahl, E. (1996). The linguistic realization of information packaging. Linguistics, 34(3), 459-519.

Weber, T., Schönhuber, M., & Grijzenhout, J. (2018). The prosody of yes/no-questions in German first language acquisition. *Conference on Phonetics & Phonology in German-speaking countries (P&P13)*,

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

Wieman, L. A. (1976). Stress patterns of early child language. *Journal of Child Language, 3*(2), 283-286.

Xu, L. J. (2004, Mar). 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. *Speech Prosody*,

Yang, X. (2002). Restrictive focus in child Mandarin. *Contemporary Linguistics, 3*, 225-237.

Zhou, P., Su, Y., Crain, S., Gao, L. Q., & Zhan, L. K. (2012, Sep). 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.