Comprehension of prosodically and syntactically marked focus in Cantonese-speaking children with and without Autism Spectrum Disorder

Article

Accepted Version

Comprehension of prosodically and syntactically marked focus in Cantonese-speaking children with and without Autism Spectrum Disorder

Haoyan GE

School of Education and Languages, Hong Kong Metropolitan University, Hong Kong SAR, China

Fang LIU

School of Psychology & Clinical Language Sciences, The University of Reading, UK

Hoi Kwan YUEN

School of Education and Languages, Hong Kong Metropolitan University, Hong Kong SAR, China

Aishu CHEN

School of English Education, Guangdong University of Foreign Studies, China

Virginia YIP

Department of Linguistics and Modern Languages & Childhood Bilingualism Research Centre, The Chinese University of Hong Kong, Hong Kong SAR, China

Correspondence

Haoyan Ge, School of Education and Languages, Hong Kong Metropolitan University, Ho Man Tin, Kowloon, Hong Kong SAR, China.

E-mail:hge@hkmu.edu.hk

Acknowledgements

The authors would like to thank the children and their families who participated in this study and gratefully acknowledge the support from the Heep Hong Society. We thank Arlene Lau and Christina Chan for their assistance with the study. The work described in this paper was substantially supported by two grants to the first author from the Hong Kong Research Grants Council – Faculty Development Scheme (Project Reference No.: UGC/FDS16/H13/19) and Katie Shu Sui Pui Charitable Trust – Research and Publication Fund (Project Reference No.:

KS2018/2.2). FL was supported by a European Research Council (ERC) Starting Grant (CAASD, 678733). We thank the anonymous reviewers for their constructive comments and helpful suggestions on earlier versions of this manuscript.

Declarations

Funding

The work described in this paper was substantially supported by two grants to HG from the Hong Kong Research Grants Council – Faculty Development Scheme (Project Reference No.: UGC/FDS16/H13/19), and Katie Shu Sui Pui Charitable Trust – Research and Publication Fund (Project Reference No.: KS2018/2.2). FL was supported by a European Research Council (ERC) Starting Grant (CAASD, 678733).

Competing interests

The authors have no competing interests to declare that are relevant to the content of this article.

Ethics approval

The study was approved by the ethics committee and conducted in accordance with the ethical standards at Hong Kong Metropolitan University.

Consent to participate

Informed consent was obtained from all adult participants and legal guardians of the children.

Author contributions

HG: Conceptualization, Funding acquisition, Methodology, Investigation, Formal analysis, Writing – Original Draft. **FL**: Conceptualization, Writing – Original Draft. **HY**: Investigation, Formal analysis, Writing – Original Draft. **AC**: Writing – Original Draft. **VY**:

Conceptualization, Writing – Original Draft. All authors contributed to the article and approved the submitted version.

Abstract

This study investigated the comprehension of prosodically and syntactically marked focus by 5- to 8-year-old Cantonese-speaking children with and without Autism Spectrum Disorder (ASD). Children listened to question-answer dialogues while looking at pictures depicting the scenarios, and judged whether the answers were correct responses to the questions. The results showed that children with ASD exhibited typically developing (TD)-like performance in the use of syntactic cues to understand focus, although they were significantly slower than their TD peers. However, children with ASD had more difficulties than their TD peers in utilizing prosodic cues in focus comprehension. These findings suggest that the comprehension difficulties found in children with ASD are domain-selective, and children with ASD are sensitive to language-specific focus marking strategies.

Keywords: Comprehension of focus, Cantonese-speaking children, Autism Spectrum Disorder, syntax, prosody

Comprehension of prosodically and syntactically marked focus in Cantonese-speaking children with and without Autism Spectrum Disorder

Autism spectrum disorder (ASD) is a neurodevelopmental disorder characterized by deficits in social interaction, along with a propensity to engage in repetitive behaviors or have restricted interests (American Psychiatric Association, 2013). The study of language acquisition in children with ASD has received considerable attention in the past decades (Boucher, 2012; Eigsti, de Marchena, Schuh, & Kelley, 2011; Hudry et al., 2010; Rapin & Dunn, 2003; Tager-Flusberg, Paul, & Lord, 2005). Previous studies suggest that children with ASD are generally impaired in their ability to comprehend linguistic cues, such as pragmatics (Kelley, Paul, Fein, & Naigles, 2006; Loukusa & Moilanen, 2009), prosody (McCann & Peppé, 2003; Paul, Augustyn, Klin, & Volkmar, 2005; Peppé, McCann, Gibbon, O'Hare, & Rutherford, 2007), and morphosyntax (Eigsti & Bennetto, 2009; Eigsti, Bennetto, & Dadlani, 2007; Roberts, Rice, & Tager-Flusberg, 2004). As far as prosody is concerned, children with ASD have difficulty with speech turning, affect, signaling phrase boundaries, and emphasizing information with contrastive accents (Paul et al., 2005; Peppé et al., 2007). In terms of morphosyntax, children with ASD fail to correctly interpret grammatical aspect (Zhou, Crain, Gao, Tang, & Jia, 2015), clitics (Terzi, Marinis, & Francis, 2016), and sentence grammaticality (Eigsti & Bennetto, 2009).

Nonetheless, recent studies on comprehension of morphosyntax and prosody in children with ASD have revealed a different picture. Specifically, the performance of children with ASD has been found to match their typically developing (TD) peers when interpreting grammatical structures including *wh*-questions (Goodwin, Fein, & Naigles, 2012), grammatical aspect (Su, & Naigles, 2021; Tovar, Fein, & Naigles, 2015), and

word order (Su & Naigles, 2019). Moreover, children with ASD show TD-like identification of statement-question intonation (Wang, Beaman, Jiang, & Liu, 2022), and can use prosody to interpret syntactic ambiguity as effectively as their TD peers (Diehl, Friedberg, Paul, & Snedeker, 2015). The previous findings raise the possibility that the comprehension ability of children with ASD may not be impaired in all linguistic domains. Given that previous comprehension studies focus on one domain of linguistic cues in children with ASD, it is still unclear how children with ASD integrate different linguistic cues in sentence comprehension.

Against this background, the current study aims to investigate how Cantonese-speaking children with ASD use both syntactic and prosodic cues to comprehend focus, compared to typically developing (TD) children and adults. Focus is a key concept of informational structure. It commonly refers to new or contrastive information in a sentence. The interpretation of focus in Cantonese involves multiple levels of linguistic knowledge, including syntax, prosody and pragmatics. The realization of focus is language-specific: Languages choose some language modules, such as grammar, prosody or morphology to encode focus. In languages with relatively strict word order like English, focus is typically realized by assigning an accent to the focal element(s) (Gussenhoven, 1983). Other languages, like Spanish and Italian, prefer syntactic means of focus-marking, i.e., to have a syntactic position where the focus is typically placed (Zimmermann & Onea, 2011). Although focus in Cantonese can be realized by syntactic or/and prosodic means (Gu & Lee, 2007; Matthews & Yip, 2011; Wu & Xu, 2010), syntactic cues are primarily used to interpret focus by Cantonese-speaking adults (Lee, 2019; Matthews & Yip, 2011). Focus provides us with an ideal testing ground to examine how Cantonese-speaking children with and without ASD integrate different

levels of linguistic information and which linguistic cue they prefer to use in sentence comprehension.

In this paper, we investigated the comprehension of focus by five to eight-year-old Cantonese-speaking children with and without ASD. The realization of focus in Cantonese is first introduced, along with reviews on previous studies regarding the acquisition of focus. Research questions and hypotheses of the study are then presented, followed by results from our comprehension experiment. Lastly, findings concerning the research questions are discussed.

Realization of focus in Cantonese

Focus commonly refers to new or contrastive information in a sentence. For instance, focus in answer (1) presents *apple* as nonpresupposed information about question (1). Focus becomes contrastive if it rejects a stated alternative in the context (Chafe, 1976). For example, the focused element *apple* in (2) forms a contrast with the alternative *pear* mentioned in question (2).

(1) Question: What did John eat?

Answer: He ate an [apple]$_F$.

(2) Question: Did John eat the pear?

Answer: No, he ate the [apple]$_F$.

It has been widely acknowledged that languages differ in their linguistic devices used to realize focus and the extent to which the same devices are used. In English, focus is typically realized by assigning prosodic prominence to the focal element(s), manifested primarily in expanded pitch range, accompanied by increased intensity and longer duration (Gussenhoven, 1983). For instance, the answer to question (1) would typically be uttered as (3a), where *APPLE* is accented (capitalization denotes

accentuation). The answer (3a) with accentuation on the object is felicitous to question (1), while (3b) with accentuation on the verb *ATE* is not.

(3) a. John ate an [APPLE]$_F$.

b. # John ATE an [apple]$_F$

Unlike English, the use of prosody to mark focus is highly constrained in Cantonese, a tonal language with six contrastive lexical tones (Chao, 1947). Specifically, there is no clear evidence for on-focus pitch expansion in Cantonese (Man, 2002; Wu & Xu, 2010). Instead, longer duration and higher intensity are manifested in Cantonese focused elements (Gu & Lee, 2007; Wu & Xu, 2010). For instance, in (4b), the subject *WU4LEI2* "fox" is accented with increased duration and intensity. Compared to English, Cantonese uses focus particles (FP) and word order to a larger extent to achieve the same purpose (Chao, 1947; Fung, 2000; Lee, 2019; Matthews & Yip, 2011). For example, the FP *hai6* could be imposed before the focused element to mark focus, as in (4a).

(4) Person A: 兔仔 食緊 紅蘿蔔?

tou3zai2 sik6gan2 hung4lo4baak6

rabbit eat-PROG carrot

"Is the rabbit eating the carrot?"

Person B: a. 唔係， 係 [狐狸]$_F$ 食緊 紅蘿蔔。

m4hai6 hai6 wu4lei2 sik6gan2 hung4lo4baak6

No FP fox eat-PROG carrot

b. 唔係， **[狐狸]$_F$** 食緊 紅蘿蔔。

m4hai6 **WU4LEI2** sik6gan2 hung4lo4baak6

No fox eat-PROG carrot

'No, the fox is eating the carrot.'

Cantonese *hai6* is not equivalent to *only* in English. *Hai6* only specifies the focused element and introduces an alternative but does not contribute to the truth conditions of the sentences, as *only* does in English. Cantonese uses other FPs, including *zing6hai6* "only", *zaa3* "only", *ze1* "only", and *zau6* "only", in different sentence positions to convey the focus meaning of *only* (Fung, 2000; Lee, 2019; Matthews & Yip, 2011).

Previous theoretical studies on Cantonese suggest that the rich inventory of FPs in Cantonese makes the use of prosody optional to encode focus meaning (Lee, 2019; Matthews & Yip, 2011). However, it is far from clear which linguistic cue (syntax or prosody) is more important for the comprehension of focus by Cantonese-speaking children with and without ASD.

Comprehension of focus in TD children

Knowledge of focus is crucial for effective communication because the interlocutors need to distinguish information provided in the preceding context, e.g., "*rabbit*" in (4), and an alternative, e.g., "*FOX*" in (4), to respond appropriately. Therefore, how children use linguistic cues to comprehend focus has been a central issue in the field of language acquisition for decades (Chen, 2010; Chen, Szendrői, Crain, & Höhle, 2019; Crain, Ni, & Conway, 1994; Höhle, Berger, Müller, Schmitz, & Weissenborn, 2009; Höhle, Fritzsche, & Müller, 2016; Paterson, Liversedge, White, Filik, & Jaz, 2006; Szendrői, Bernard, Berger, Gervain, & Höhle, 2018; Zhou & Crain, 2010; Zhou, Su, Crain, Gao, & Zhan, 2012).

Previous research on the comprehension of focus in TD children has yielded controversial results. Some researchers suggest that children's interpretation of focus with *only* does not seem adult-like until school age. In a picture-verification task, Crain et al. (1994) asked 3- to 6-year-old English-speaking children to judge whether

sentences like *Only the cat is holding a flag* and *The cat is only holding a flag* were accurate descriptions of a picture in which a cat was holding a flag, a duck was holding a flag and a balloon, and a frog was holding a balloon. They found that English-speaking children associated *only* with the verb phrase regardless of its position in the sentence, different from the adult controls. Using a similar task, Paterson et al. (2006) investigated the interpretation of focus in *only*-sentences by English-speaking children at an elder age (7;0–10;0). Their results showed that children interpreted focus with and without *only* in the same way. Non-adult-like patterns were also observed in Mandarin-speaking children. Chen (1998) used a sentence-correction task to assess the use of different syntactic and prosodic cues by 5- to 13-year-old Mandarin-speaking children in determining given/new information. Results showed that Mandarin-speaking children rely more heavily on prosodic information than on word order in determining given and new information, whereas Mandarin-speaking adults considered word order to be the major cue for given/new information. Zhou and Crain (2010) asked 4-year-old Mandarin-speaking children to interpret focus marked by FPs like *Zhiyou Yuehan chi-le pingguo* 'Only John ate an apple' and *Shi Yuehan chi-de pingguo* 'It is John who ate an apple'. It is found that Mandarin-speaking 4-year-olds tend to associate FPs *zhiyou* 'only' and *shi* 'be' with the verb phrase, whereas adults uniquely associate them with the subject.

Apart from the investigations on the syntactically marked focus, previous studies also found that children before age 8 have not achieved an adult-like understanding of prosodically marked focus with *only*. Using a truth-value judgement task, Gualmini, Maciukaite and Crain (2003) investigated whether 4- to 5-year-old English-speaking children were sensitive to prosodic information in comprehending sentences with preverbal *only* in two conditions: prosodic prominence placed either on the indirect

object (e.g., *The Troll only brought an onion ring to SUPERMAN*) or on the direct object (e.g., *The Troll only brought an ONION RING to superman*). The results showed that children interpreted the sentences in the two conditions in the same way, suggesting that they were unable to use prosodic information alone to correctly interpret focus in *only*-sentences. In a similar vein, studies on German-speaking children (5;1–7;8) (Hüttner, Drenhaus, van der Vijver, & Weissenborn, 2004) and Dutch-speaking children (4;1–6;10) (Szendrői, 2004) also showed children's failure to understand prosodically marked focus in sentences with FPs in an adult-like fashion.

However, whether the comprehension of syntactically and prosodically marked focus is problematic for children has been challenged. Some researchers argued that adult-like comprehension of focus is possible. In an eye-tracking study, Höhle et al. (2009) asked German-speaking 2- to 4-year-old children to listen to sentences with either accented *auch* "also" (*Toby hat AUCH eine Puppe* "Toby has ALSO a doll") or unaccented *auch* "also" (*Toby hat auch eine ENTE* "Toby has also a DUCK") and look at a picture in which one child Toby had a doll and a duck and the other child had only one doll. The fixation patterns suggested that children could use prosodic information from at least 3 years of age onwards to identify the correct focus domain and relate focus with its corresponding alternative set. Höhle et al. (2016) further investigated the comprehension of focus with *nur* "only" in German-speaking 4-year-old children. Although the offline judgement data replicated previous findings of non-adult-like comprehension of focus, children's eye gaze data revealed adult-like patterns of focus interpretation. In a similar eye-tracking study, Zhou et al. (2012) investigated the use of prosodic information by 5-year-old Mandarin-speaking children in online comprehension of focus with *zhiyou* 'only'. Their results also showed adult-like patterns in the eye-tracking task but non-adult-like explicit judgements. The high

variation in children's performance across studies is unclear but could be attributed to the nature of the tasks, which may prevent children from putting their knowledge into full use.

Apart from task differences, other researchers suggested that children's failure in the comprehension of focus in sentences with *only* is not necessarily due to their lack of knowledge of focus but due to extra demands from *only*. Therefore, Chen (2010) examined the processing of accentuation as a cue to comprehend focus in simple subject-verb-object (SVO) sentences without *only* by 4- to 5-year-old Dutch-speaking children. The results show that Dutch-speaking children can use prosodic information to interpret focus in an adult-like manner, although their response times were longer than adults. Szendrői et al. (2018) also reveal an adult-like interpretation of prosodically marked subject and object focus in 3- to 6-year-old English-, German- and Dutch-speaking children. Furthermore, their study showed that children of different languages had acquired cross-linguistic differences in the use of focus markers in comprehension very early on. In a recent study, Chen et al. (2019) presented Mandarin-speaking children (aged 3-5) with simple SVO sentences without *only* and asked them to judge whether the sentence matched the picture or not. Different from Chen (1998), their results showed that Mandarin-speaking children preferred syntactic information over prosodic information to interpret focus, indicating that children as early as 3 years of age were attuned to specific means of focus marking.

Based on the previous findings, we noticed that firstly, it is not yet fully clear whether TD children at the beginning of school-age have an adult-like comprehension of focus. Secondly, most of the previous studies on TD children examined focus comprehension involving *only*. This is potentially questionable because these tasks also test children's ability to comprehend *only* rather than focus itself, which may increase

the demands on children and thus influence the results. Lastly, previous studies showed that TD children exhibit language-specific abilities in understanding focus across languages. However, whether the results can be generalized to children with ASD is still unclear.

Comprehension of focus in children with ASD

Previous studies on children with ASD mainly concentrated on the comprehension of prosodic focus using a test of prosodic abilities, the Profiling Elements of Prosodic Systems – Children (PEPS-C; Peppé & McCann, 2003). The PEPS-C includes testing of four communication areas where prosody plays a crucial role: interaction, affect, chunking and focus. Regarding the comprehension of focus is concerned, the PEPS-C measures the ability to perceive contrastive focus accents. In this task, children see two colors on the screen and hear sentences with focus accents on different colors (e.g., *I wanted BLUE and black socks* vs *I wanted blue and BLACK socks*). Then they need to point to the color that was focused. Peppé et al. (2007) found that English-speaking children with ASD aged 6-13 years made significantly more errors, compared to the matched TD group. On the other hand, there are studies that children with ASD between the age of 6 and 16 achieve TD-like interpretation of focus (Filipe, Frota, & Vicente, 2018; Järvinen-Pasley, Peppé, King-Smith, & Heaton, 2008).

To the best of our knowledge, only one pilot study has investigated the comprehension of prosodically marked focus in Cantonese-speaking children with ASD (Zhou et al., 2021). In their study, eighteen 7- to 9-year-old Cantonese-speaking children with and without ASD participated in a naturalness rating task. Children first heard question-answer dialogues and viewed corresponding pictures. The questions generated a broad focus (e.g., *What's happening?*), whereas the answers presented with

either neutral prosodic prominence (e.g., *Mr. Zhang is flying the plane*) or prosodic prominence on the subject (e.g., *MR. ZHANG is flying the plane*), giving rise to matched and mismatched conditions. Then, children were asked to rate the naturalness of the dialogues on a scale from 1 (very unnatural) to 5 (very natural). Their results showed that Cantonese-speaking children with and without ASD tended to rate the dialogues as 5 (very natural) regardless of the conditions, suggesting that both groups of children were not able to interpret the prosodic focus. However, their findings might be masked by the nature of the task. Different from an identification/verification task, a naturalness rating provides less restricted choices which increase the choice difficulty and thus has been seldom used in children. It is possible that the children might be able to correctly interpret prosodic focus but failed to understand the task. Furthermore, Cantonese-speaking children might rely heavily on morphosyntactic means other than prosodic cues to interpret focus, especially given the focus marking strategies in Cantonese.

Taken together, it is far from clear whether TD-like comprehension of focus is possible in children with ASD. The previous findings also highlight the importance of understanding language-specific patterns in children with ASD. It is thus crucial to include both syntactically and prosodically marked focus to examine whether children with ASD are subject to language-specific strategies.

The current study

The current study investigated the comprehension of both syntactically and prosodically marked focus in simple SVO sentences without *only* by 5- to 8-year-old Cantonese-speaking children with and without ASD, whose age, receptive language ability, nonverbal intelligence, and working memory were measured and controlled for. We focus on this age population not only because they are under-studied, but also

because they provide a wealth of information about how focus is interpreted at the beginning of school age. We raise three research questions. First, could Cantonese-speaking TD children have adult-like comprehension of focus? As noted above, the findings of the prior studies are mixed. As our study examined focus in sentences without *only*, it would require relatively fewer demands on children compared to the interpretation of focus in *only*-sentences. Based on the previous findings on comprehension of focus in simple sentences by TD children (Chen, 2010; Chen et al., 2019; Szendrői et al., 2018), we predicted that Cantonese-speaking TD children would exhibit adult-like comprehension of focus.

Second, would Cantonese-speaking children with ASD show TD-like comprehension of focus? As the previous study tentatively suggests that the interpretation of prosodically marked focus is difficult for Cantonese-speaking children with ASD around 7 to 9 years of age (Zhou et al., 2021), we hypothesized that children with ASD at a younger age (5 to 8 years of age) in this study would also have comprehension difficulties and not able to interpret focus in a TD-like fashion.

Third, which linguistic cue (syntax or prosody) would Cantonese-speaking children with and without ASD prefer in interpreting focus? The previous cross-linguistic studies showed that language-specific marking of focus is acquired as early as 3 years of age (Chen et al., 2019; Szendrői et al., 2018), we predicted that both Cantonese-speaking children with and without ASD shall have acquired the language-specific marking of focus by the age of 8. Specifically, the children in the current study would rely more on syntactic cues more than prosodic cues in interpreting focus, similar to Mandarin-speaking children (Chen et al., 2019), or they could correctly interpret syntactically marked focus but were not sensitive to prosodically marked focus.

Method

Participants

Forty-four Cantonese-speaking children with high-functioning ASD and fifty-four TD children participated in this study. Twenty-eight Cantonese-speaking adults served as the comparison groups. All participants were born in Hong Kong and acquired Cantonese as their first language. TD children had no family history of diagnosed developmental disorders or impairments. Children with ASD were mainly recruited from the Heep Hong Society, a leading education and rehabilitation organization in Hong Kong that offers diverse support services to children with special education needs. Parents' informed consent was obtained prior to the study. All procedures were approved in accordance with the research ethical committee at the university where the testing took place. Demographic information of the three groups of participants is shown in Table 1. TD children and children with ASD were matched on age, receptive vocabulary, non-verbal IQ, and working memory. The Cantonese Receptive Vocabulary Test (CRVT; Cheung et al., 1997) was used to assess the children's receptive Cantonese vocabulary knowledge. Non-verbal IQs were assessed with the Primary Test of Nonverbal Intelligence (PTONI; Ehrler & McGhee, 2008) by the third author, a clinical psychologist. Nonverbal IQs for children with ASD ranged from 68 to 143 (mean = 111.61), while the IQs for TD children ranged from 71 to 144 (mean = 113.53). The Backward Digit Span task, based on the procedure included in the McCarthy Scales of Children's Abilities (McCarthy, 1972), was used to evaluate participants' working memory. The task was composed of two sections, in which the instructors said a list of numbers, and the participant was asked to recite the numbers backwards. The numbers increased by one item for every correct response, and the largest set consisted of six numbers. If the participant answered incorrectly twice in a

row, the instructor moved on to the next section (Nesbitt, Baker-Ward, & Willoughby, 2013). A score was assigned based on the largest, correctly reported set, with a maximum score of 10. T-tests revealed no group difference between TD children and children with ASD in terms of PTONI ($t(90) = 0.260$, $p = 0.398$), CRVT ($t(90) = 0.241$, $p = 0.810$) and working memory ($t(90) = -0.487$, $p = 0.628$).

Autism or autism spectrum diagnoses were validated with the Autism Diagnostic Observation Schedule™, Second Edition (ADOS™-2; Lord et al., 2012) by the first author who has the qualification of administering and coding the ADOS-2 for clinical and research reliability. As a semi-structured standardized assessment tool, the ADOS-2 is used to measure social and communication behaviours which contribute to a diagnosis of autism. In the current study, Module 3 was selected based on children's language and developmental levels. Children were classified as ASD when they received a total score of ≥ 7 (N = 26), and they were classified as autism if their total scores were ≥ 9 (N = 18).

Design and materials

We adopted Chen's (2010) RT paradigm to examine children's focus comprehension. In each trial, participants were presented with a question-answer dialogue between a pig and a robot. They were then asked to judge whether the answer from the robot made sense as a response to the question from the pig. A 3 × 2 × 2 design was used to manipulate *Group* (ASD, TD, adults), *Focus Marking* (FP, prosody) and *Context* (match, mismatch). For the experimental dialogues, there were one version of each question from the pig and four versions of each answer from the robot. Each question from the pig consisted two parts: the first part (e.g., *Who is eating the carrot?*) triggered new information to be expected on subject, and the second part (e.g., *I think the monkey*

is eating the carrot) elicited subject focus in the comment. The variables *Focus Marking* (FP, prosody) and *Context* (match, mismatch) were embedded in the answers from the robot, which gave rise to four experimental conditions: (a) focus marked with FP in a matched context, (b) focus marked with prosody in a matched context, (c) focus marked with FP in a mismatched context, and (d) focus marked with prosody in a mismatched context. Examples of the four conditions are illustrated in Table 2, where the focused words are bracketed and prosodic prominence is underlined. The experimental auditory stimuli in Table 2 correspond to the picture stimuli in Figure 1, in which two cartoon characters (e.g., *fox* vs. *monkey*) are performing the same action (e.g., *eating*) to two different objects (e.g., *cabbage* vs. *carrot*).

The auditory stimuli were recorded by a female native speaker of Cantonese at 44.1 kHz sampling frequency with 16 bits resolution. She was asked to produce the stimuli with appropriate or inappropriate FP and prosody for four experimental conditions. Each stimulus was scaled to 70 dB SPL in mean intensity using Praat (Version 6.0.39; Boersma & Weenink, 2018). To ensure that prosody was placed in the right position, the answer sentences of the experimental dialogues were subjected to a phonetic analysis using Praat. All stimuli were cross-checked by two native speakers of Cantonese (one male and one female) to ensure the stimuli were natural. To measure the reliability of the stimuli, we carried out a reliability analysis on the two native speakers' judgements, using the *psych* package (Revelle, 2019). A good measurement of reliability (α Cronbach = 0.84) was observed. In total, 80 experimental dialogues and 40 fillers were distributed over the four conditions via a Latin Square design. Four lists of dialogues were created such that each dialogue appeared in every experimental condition but not in the same list. Each participant was presented with only one list that

included 32 dialogues (4 experimental conditions × 5 experimental dialogues + 10 fillers + 2 practice items).

Procedure

Each testing session began with two practice trials to familiarize the participants with the experiment. Each trial was set up in E-Prime 3.0 as follows. First, a cross appeared in the center of the screen. Then, a question-answer dialogue between the pig and the robot was played right after, with a 2000-ms interval between the question and the answer. The two options, "YES or NO," were displayed on the screen at the end of the answer. The participants were instructed to rest their thumbs on an RT box and press the button to indicate their response as quickly as possible, but not before the end of the answer sentence. If the answer was a correct response to the question, they were asked to press the "YES" button (on the left side of the RT box), otherwise the "NO" button (on the right side of the RT box). Therefore, the "YES" responses to the questions were not truly "correct", but those that the children "thought" were correct. For example, the answer to Condition (d) where focus was marked by prosody in a mismatched context was an incorrect response to the question in Table 2. However, if a child thought the answer was correct without realizing the inappropriate placement of prosodic prominence on the unfocused element, he or she would give a "YES" response. "YES-NO" judgments and RTs were recorded at the end of each answer sentence until a button was pressed using E-prime. The participants could take two breaks of any length in the middle of the task. It took each participant 15 – 25 minutes to complete the experiment. The participants were unaware of the purpose of the experiment and received cash coupons as compensation.

Data analysis

Only RTs in the experimental trials where the answers were judged as "Correct" were included for further analysis. Raw RTs smaller than 200 ms or above 2.5SD were further excluded. We conducted the Shapiro–Wilk test on the remaining raw RTs in the R statistical program (R Core Team, 2022) to examine its normality. As the RTs were not normally distributed ($W= 0.765$, $p < 0.001$), we log-transformed the RT data to reduce the non-normality of residuals. To measure the task reliability, we carried out a reliability analysis on the log-transformed RTs comprising 80 items in the R statistical program, using the *psych* package (Revelle, 2019). Cronbach's alpha showed that the task reached acceptable reliability ($\alpha = 0.881$). To examine how *Group*, *Focus marking*, and *Context* affect participants' comprehension of focus, we used linear mixed-effects models in the *lme4* package (Bates, Mächler, Bolker, & Walker, 2015) for all analyses in R. In the models, we included fixed factors of *Group* (Adults, ASD, TD), *Focus Marking* (FP, prosody), and *Context* (match, mismatch) with *Participant* and *Item* as random factors. Given the unbalanced gender distribution between the two groups of children, *Gender* was also included as a factor in the models for statistical analysis. Since the factor of *Gender* did not lead to significant differences in children's performance in the tasks, it was not included in the model selection. Thus, we assume that female participants performed similarly to male participants in the current study. The dependent variables were "YES–NO" judgment (YES=1, NO=0) and log-transformed RTs. Effects were tested for significance by model comparison. To assess the goodness of the models, we compared the models using the $\chi 2$- distributed likelihood ratio and its associated p-value. The model with a smaller Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) was considered as a better fit (Baayen, Davidson, & Bates, 2008). Significant interaction

effects between fixed effects were followed by pairwise comparisons with "tukey" adjustment for multiple comparisons using *emmeans* package (Lenth, 2018).

Results

Yes-No judgements

Figure 2 presents the mean percentage of "YES" responses in Cantonese-speaking adults, children with ASD and TD children in four conditions. Table 3 shows the significance of factors in the final model. There are significant main effects for each factor, a significant three-way interaction between *Group*, *Focus Marking* and *Context*, and three two-way interactions involving each pairing of these three factors.

Regarding Cantonese-speaking adults, post-hoc analyses show that their proportion of "YES" responses was significantly higher in matched contexts than in the mismatched contexts in both FP marked focus (*Estimate* = 0.357, *SE* = 0.028, *t* = 12.775, $p < 0.001$) and prosodically marked focus (*Estimate* = 0.136, *SE* = 0.028, *t* = 4.854, $p < 0.001$). Moreover, they gave significantly less "YES" responses in comprehending FP marked focus than prosodically marked focus in mismatched contexts (*Estimate* = -0.193, *SE* = 0.028, *t* = -6.898, $p < 0.001$). TD children gave significantly more "YES" responses to FP marked focus in matched contexts than in mismatched contexts (*Estimate* = 0.085, *SE* = 0.020, *t* = 4.231, *p* = 0.0014), but not to prosodically marked focus across the contexts (*Estimate* = 0.015, *SE* = 0.020, *t* = 0.736, *p* = 0.999, n.s.). Similar to TD children, the proportion of "YES" response of children with ASD to FP marked focus was significantly higher in matched contexts than in mismatched contexts (*Estimate* = 0.059, *SE* = 0.023, *t* = 3.532, *p* = 0.032), whereas their response to prosodically marked focus was similar across the two contexts (*Estimate* = 0.024, *SE* = 0.023, *t* = 1.056, *p* = 0.996, n.s.).

Post-hoc pairwise comparison showed that group difference lay crucially between adults and TD children (*Estimate* = -0.125, *SE* = 0.028, *t* = -4.458, *p* <0.001), as well as between adults and children with ASD (*Estimate* = -0.093, *SE* = 0.030, *t* = -3.150, *p* = 0.006), whereas TD children and children with ASD performed comparably (*Estimate* = -0.032, *SE* = 0.024, *t* = -1.345, *p* = 0.373, n.s.). Specifically, there was no significant difference among adults, TD children and children with ASD in matched contexts, regardless of the focus marking strategy. When comprehending FP marked focus in mismatched contexts, adults gave significantly less "YES" response than TD children (*Estimate* = -0.281, *SE* = 0.035, *t* = -7.998, *p* < 0.001) and children with ASD (*Estimate* = -0.279, *SE* = 0.037, *t* = -7.575, *p* < 0.001), whereas there was no significant difference between TD children and children with ASD (*Estimate* = -0.001, *SE* = 0.031, *t* = -0.046, *p* = 1.000, n.s.). In terms of prosodically marked focus in mismatched contexts, adults' proportion of "YES" response was significantly lower than TD children (*Estimate* = -0.166, *SE* = 0.035, *t* = -4.722, *p* < 0.001) and children with ASD (*Estimate* = -0.111, *SE* = 0.037, *t* = -3.010, *p* < 0.001). Again, TD children and children with ASD performed similarly in comprehending prosodically marked focus in mismatched contexts (*Estimate* = -0.055, *SE* = 0.031, *t* = -1.796, *p* = 0.820, n.s.).

Reaction times

The mean RTs and log-transformed RTs of the three groups in four conditions are shown in Table 4 and Figure 3 respectively. Recall that a linear mixed-effects model was applied to participants' log-transformed RTs to examine the effects of *Group*, *Focus Marking*, *Context*, and their interactions. The final model included the effects of *Group*, *Focus Marking*, *Context* and their interactions, which is given in Table 5.

There was a main effect of *Group*, a two-way *Group* × *Context* interaction and a two-way *Focus Marking* × *Context* interaction. Post-hoc analyses showed that the group differences mainly lay between children with ASD and adults (*Estimate* = -0.260, *SE* = 0.044, *t* = -5.934, $p < 0.001$) and between children with ASD and TD children (*Estimate* = -0.170, *SE* = 0.035, *t* = 4.847, $p < 0.001$), suggesting that both adults and TD children were much faster than children with ASD in understanding focus across conditions, regardless of focus marking and contexts. Post-hoc pairwise comparisons also showed that adults were significantly faster than TD children in comprehending FP marked focus in matched contexts (*Estimate* = -0.158, *SE* = 0.047, *t* = -3.388, *p* = 0.039), whereas they performed comparably in mismatched contexts (*Estimate* = -0.068, *SE* = 0.049, *t* = -1.382, *p* = 0.966, n.s.). In comprehending prosodically marked focus, adults and TD children showed similar performance in matched (*Estimate* = -0.112, *SE* = 0.047, *t* = -2.386, *p* = 0.420, n.s.) and mismatched contexts (*Estimate* = -0.022, *SE* = 0.048, *t* = -0.459, *p* = 1.000, n.s.).

For Cantonese-speaking adults, they were significantly much slower in comprehending both FP marked focus (*Estimate* = -0.112, *SE* = 0.031, *t* = -3.639, *p* = 0.015) and prosodically marked focus (*Estimate* = -0.112, *SE* = 0.029, *t* = -3.035, *p* = 0.037) in mismatched contexts than in matched contexts. Regarding TD children, they showed similar RTs across the four conditions, showing that focus marking and context play a little role in comprehension. While children with ASD showed similar RTs to FP marked focus across the two contexts, they were significantly faster in responding to prosodically marked focus in mismatched contexts than in matched contexts (*Estimate* = 0.059, *SE* = 0.033, *t* = 2.811, *p* = 0.003).

Discussion

The current study investigated how 5 to 8-year-old Cantonese-speaking children with and without ASD use syntactic and prosodic information to comprehend focus. Our results indicate that Cantonese-speaking TD children were able to use syntactic cues to interpret focus, as they gave significantly less "YES" response to FP marked focus in mismatched contexts than in matched contexts, although to a lesser extent than Cantonese-speaking adults. In terms of prosodically marked focus, adults were more accurate in both matched and mismatched contexts compared to TD children, showing that TD children were not sensitive to prosodically marked focus regardless of the context. TD children seem to have adult-like comprehension of syntactically marked focus, but their interpretation of prosodic focus is not adult-like. We found that, in a task where children did not have to interpret semantic operators, such as *only*, they were able to show adult-like understanding of syntactically marked focus, consistent with our hypothesis and previous studies on TD children's comprehension of focus in simple sentences (Chen, 2010; Chen et al., 2019; Szendrői et al., 2018). While the primary goal of the present study was to understand the comprehension of focus in ASD, this study also provides novel information about the possibility of adult-like focus comprehension in Cantonese-speaking TD children.

In terms of the second research question that whether children with ASD have TD-like comprehension, we noticed two patterns. First, children with ASD were as good as TD children in making use of syntactic cues in interpreting focus, even though they were slower. Second, children with ASD were faster when prosody and focus were mismatched than when prosody and focus were matched, while TD children did not show the same pattern. Our results provide new empirical evidence that Cantonese-speaking children with ASD can achieve TD-like performance in language

development, as their comprehension patterns of syntactically marked focus is similar to TD children. Meanwhile, we should also note that children with ASD exhibit non-TD-like use of prosodic information in understanding focus, even after controlling for the effects of age, receptive vocabulary, nonverbal intelligence, and working memory. The findings suggest that the comprehension difficulty observed in children with ASD is rather domain-specific, and TD-like performance in some linguistic domains is possible.

Regarding the third research question, Cantonese-speaking adults showed a stronger preference in utilizing syntactic cues over prosodic information, although they make use of both syntactic and prosodic information to comprehend focus. This finding provides further empirical evidence supporting previous results on Cantonese focus realization: Cantonese relies heavily on syntax for focus marking, in contrast to English that rests on prosody to realize focus (Fung, 2000; Lee, 2019). This language-specific strategy of focus marking is also observed in Cantonese-speaking children with and without ASD. By age 8, both groups of children have developed a language-specific preference, i.e., syntactic marking, to interpret focus in Cantonese, whereas their ability to make use of prosodic cues in focus comprehension has not yet fully developed. Our findings are consistent with previous results that TD children attune to the specific means of focus marking from an early age (Chen et al., 2019; Szendrői et al., 2018). The current findings further confirm that language-specific marking of focus can be acquired at the beginning of school age, from the perspective of typical and atypical language development in children with and without ASD. This language-specific strategy might enable Cantonese-speaking children with and without ASD to exhibit adult-like performance in comprehending syntactically marked focus, but give rise to developmental delays in understanding prosodically marked focus in Cantonese.

We now consider how to account for the differences between children with ASD and TD children. The first possibility could be that children with ASD have more difficulties than their TD peers with the integration of information. The comprehension of focus is not an easy task to accomplish. Children first need to identify relevant syntactic/prosodic information from multiple sources and then integrate the information with focused constituents through the context in order to accurately interpret the meaning. In particular, Cantonese-speaking children with ASD may correctly perceive syntactic or prosodic cues in sentences but struggle to integrate these into their processing of focus meaning through the question-answer dialogues. This may explain why children with ASD were slower than TD children in general when interpreting both syntactically and prosodically marked focus. This line of interpretation is also consistent with the weak central coherence hypothesis' claim of a domain-general difficulty integrating multiple information (Happé & Frith, 2006). Our findings suggest that weak central coherence may apply not only to visual and auditory processing but also to language processing in individuals with ASD, in line with the previous studies on the resolution of ambiguous words (Jolliffe & Baron-Cohen, 1999, 2000).

Nonetheless, we also found that children with ASD, compared to TD children, were faster in responding to prosodic focus in mismatched contexts than in matched contexts. This pattern was absent in their responses to syntactically marked focus, which cannot be fully explained by weak central coherence. We think this non-TD-like performance of children with ASD can be attributed to their difficulty in correctly perceiving prosodic prominence in the first place. This explanation is not unlikely, considering previous studies on the perception of prosody in children with ASD (e.g., McCann & Peppé, 2003; Paul, et al., 2005; Peppé et al., 2007). Children with ASD might take the inappropriate prosody as the correct form, reflected in their faster

response to prosodic focus in mismatched contexts, and thus fail to map the correct prosodic information to focus.

A further remark concerns the effect of task difficulty on the performance of adults and children. One anonymous reviewer pointed out that all three groups of participants showed a high number of "YES" responses in the mismatched contexts. If it was the task difficulty that lead to a high number of "YES" responses, we would expect all three groups to show similar performance across the conditions. However, we have observed a significant difference between the matched and mismatched context in comprehending FP marked focus across groups. Thus, it is unlikely that the task difficulty could explain the high proportion of "YES" responses in prosodically marked focus.

Our findings complement and extend the previous research in a number of ways. First, our study contributes to a better understanding of focus comprehension in both children with ASD and TD children at the beginning of school age. Our findings also provide new cross-linguistic empirical evidence for comprehension of focus in typical and atypical development, from the perspective of Cantonese. Second, our study advances our understanding of multiple information integration in language comprehension of children with ASD. Our findings indicate that the integration of multiple information does pose difficulty to children with ASD, consistent with the weak central coherence hypothesis. Meanwhile, our results also highlight the importance of considering other factors, such as deficits in prosodic perception, to fully account for the comprehension difficulties found in children with ASD. Teasing apart the weak central coherence and prosodic deficits in ASD will require improving our experimental design. In particular, it is important to assess weak central coherence in a variety of linguistic domains in children with ASD, which will be valuable in

understanding the language deficits associated with autism. Finally, our study is one of the very few ones that compare the use of syntactic and prosodic information in the comprehension of children with ASD. The findings also provide evidence that Cantonese-speaking children with ASD, like TD children, are sensitive to language-specific focus marking before age 8. The language-specific strategies might pose greater demands on children with ASD than their TD peers in acquiring a less preferred processing strategy in their language, such as the use of prosody to understand focus in Cantonese. It would be interesting to investigate whether and when Cantonese-speaking children with and without ASD start to adhere to prosodic information for focus interpretation.

Our findings have implications for evidence-based assessments and practice in atypical language development of children with ASD. First, our results suggest that children with ASD may not always experience more difficulties than their TD peers in all aspects of language development. Assessing their linguistic knowledge in different domains is crucial to obtain a more comprehensive picture of their language development. Second, our study highlights the importance of understanding language-specific patterns in designing future language assessment tools to detect communication problems in children with ASD. In particular, identifying cross-linguistic differences and considering language backgrounds would be helpful in designing intervention plans for children with ASD.

The current study is not without limitations. First, we focused solely on high-functioning ASD children with strong language skills. Therefore, the results may not generalize to the broader population of children with ASD. Further research is needed to extend to relatively low-functioning children with ASD. Second, our study only examined the comprehension of focus. It is still unclear how Cantonese-speaking

children with ASD would perform in the production of focus. Previous studies found that English-speaking children with ASD made more errors in focus production than TD children (Diehl & Paul, 2013; Paul et al., 2005; Peppé et al., 2007). Further research is needed to obtain a richer understanding of both comprehension and production of focus. Furthermore, the current study only examined focus in subject position, without comparing focus in verb or object positions. Further studies are desired to investigate whether and to what extent focus position would influence the comprehension of children with ASD.

Conclusion

This study is the first to examine the comprehension of prosodically and syntactically marked focus in Cantonese-speaking children with ASD. We have identified their use of prosodic and syntactic means, while comparing them to Cantonese-speaking adults and TD children. On the one hand, the similar patterns between children with ASD and TD children provide supporting evidence for TD-like syntactic abilities in the comprehension of focus by children with ASD. On the other hand, the results suggest that children with ASD might experience more difficulties in utilizing prosodic information to understand focus, relative to their TD peers. Our results also indicate that Cantonese-speaking children with ASD have acquired language-specific focus marking strategies by age 8. This study demonstrates how linguistic domains (e.g., prosody and syntax) and language-specific strategies can influence comprehension outcomes in children with ASD. The findings have implications for future assessments and evidence-based interventions for children with ASD in general.

Declarations

Ethics approval

The study was approved by the ethics committee and conducted in accordance with the ethical standards at [removed for review].

Funding

[removed for review]

Competing interests

The authors have no competing interests to declare that are relevant to the content of this article.

Consent to participate

Informed consent was obtained from all adult participants and legal guardians of the children.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

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., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

Boersma, P., & Weenink, D. (2018). Praat: Doing phonetics by computer [Computer software]. Version 6.0.39. http://www.praat.org/

Boucher, J. (2012). Research review: Structural language in autism spectrum disorder – characteristics and causes. *Journal of Child Psychology and Psychiatry*, *53*, 219–233.

Chafe, W. (1976). Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In C. N. Li (Ed.), *Subject and topic* (pp. 225–255). New York: Academic Press.

Chao, Y. R. (1947). *Cantonese Primer.* Cambridge, Massachusetts: Harvard University Press.

Chen, A. (2010). Is there really an asymmetry in the acquisition of the focus-to-accentuation mapping?. *Lingua*, *120*(8), 1926-1939.

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). Surface cues and the development of given/new interpretation. *Applied Psycholinguistics*, *19*(4), 553–582.

Cheung, P. S., Lee, K. Y., & Lee, L. W. (1997). The development of the 'Cantonese Receptive Vocabulary Test' for children aged 2-6 in Hong Kong. European *Journal of Disorders of Communication*, *32*(1), 127-138.

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

Diehl, J. J., & Paul, R. (2013). Acoustic and perceptual measurements of prosody production on the profiling elements of prosodic systems in children by children with autism spectrum disorders. *Applied Psycholinguistics*, *34*(1), 135-161.

Diehl, J. J., Friedberg, C., Paul, R., & Snedeker, J. (2015). The use of prosody during syntactic processing in children and adolescents with autism spectrum disorders. *Development and Psychopathology*, *27*(3), 867–884.

Ehrler, D. J., & McGhee, R. L. (2008). *PTONI: Primary test of nonverbal intelligence*. Austin, TX: Pro-Ed.

Eigsti, I. M., & Bennetto, L. (2009). Grammaticality judgments in autism: Deviance or delay. *Journal of Child Language*, *36*(5), 999-1021.

Eigsti, I. M., Bennetto, L., & Dadlani, M. B. (2007). Beyond pragmatics: Morphosyntactic development in autism. *Journal of Autism and Developmental Disorders*, *37*, 1007–1023.

Eigsti, I. M., de Marchena, A. B., Schuh, J. M., & Kelley, E. (2011). Language acquisition in autism spectrum disorders: A developmental review. *Research in Autism Spectrum Disorders*, *5*(2), 681–691.

Filipe, M. G., Frota, S., & Vicente, S. G. (2018). Executive functions and prosodic abilities in children with high-functioning autism. *Frontiers in Psychology*, 9(MAR), 1–10.

Fung, R. S. Y. (2000). *Final particles in standard Cantonese: semantic extension and pragmatic inference*. Doctoral *dissertation*, The Ohio State University.

Goodwin, A., Fein, D. & Naigles, L. R. (2012). Comprehension of wh-questions precedes their production in typical development and autism spectrum disorders. *Autism Research*, *5*, 109–23.

Gu, W. & Lee, T. (2007). Effects of tonal context and focus on Cantonese F0. *Proceedings of 16th International Conference Phonetic Science*, Saarbrucken, 1033-1036.

Gualmini, A., Maciukaite, S., & Crain, S. (2003). Children's insensitivity to contrastive stress in sentences with only. *University of Pennsylvania Working Papers in Linguistics*, *8*(1), 8.

Gussenhoven, C. (1983). *On the grammar and semantics of sentence accents*. Dordrecht: Foris.

Happé, F.G.E., & Frith, U. (2006). The weak coherence account: Detail-focused cognitive style in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *36*(1), 5–25.

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

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.

Hudry, K., Leadbitter, K., Temple, K., Slonims, V., McConachie, H., Aldred, C., & Charman, T. (2010). Preschoolers with autism show greater impairment in

receptive compared with expressive language abilities. *International Journal of Language & Communication Disorders*, *45*, 681–690.

Hüttner, T., Drenhaus, H., van der Vijver, R., & Weissenborn, J. (2004, June). The acquisition of the German focus particle auch 'too': comprehension does not always precede production. In *Proceedings of the 28th Annual Boston University Conference on Language Development, Online Supplement*.

Järvinen-Pasley, A., Peppé, S., King-Smith, G., & Heaton, P. (2008). The relationship between form and function level receptive prosodic abilities in autism. *Journal of Autism and Developmental Disorders*, *38*(7), 1328–1340.

Jolliffe, T., & Baron-Cohen, S. (1999). A test of central coherence theory: Linguistic processing in high-functioning adults with autism or Asperger syndrome: Is local coherence impaired? *Cognition*, *71*(2), 149–185.

Jolliffe, T., & Baron-Cohen, S. (2000). Linguistic processing in high-functioning adults with autism or asperger's syndrome. Is global coherence impaired? *Psychological Medicine*, *30*(5), 1169–1187.

Kelley, E., Paul, J. J., Fein, D., & Naigles, L. R. (2006). Residual language deficits in optimal outcome children with a history of autism. *Journal of Autism and Developmental Disorders*, *36*, 807–828.

Lee, P. P. L. (2019). *Focus manifestation in Mandarin Chinese and Cantonese: A comparative perspective.* Routledge.

Lenth, R. (2018). *emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.1.3*.

Lord, C., Rutter, M., DiLavorce, P. C., Risi, S., Gotham, K., & Bishop, S. (2012). *Autism Diagnostic Observation Schedule, Second Edition (ADOS-2) Manual (Part I): Modules 1-4*. Torrance, CA: Western Psychological Services.

Loukusa, S., & Moilanen, I. (2009). Pragmatic inference abilities in individuals with Asperger syndrome or high-functioning autism. A review. *Research in Autism Spectrum Disorders*, *3*, 890–904.

Man, V. C. (2002). Focus effects on Cantonese tones: An acoustic study. *Proceedings of 1st International Conference on Speech Prosody*, Aix-en-Provence, France, 467-470.

Matthews, S., & Yip, V. (2011). *Cantonese: A comprehensive grammar.* London, England: Routledge.

McCann, J., & Peppe, S. (2003). Prosody in autism spectrum disorders: A critical review. *International Journal of Language & Communication Disorders*, *38*(4), 325–350.

McCarthy, D. (1972). *McCarthy Scales of Children's Abilities*. New York: The Psychological Corp.

Nesbitt, K. T., Baker-Ward, L., & Willoughby, M. T. (2013). Executive function mediates socio-economic and racial differences in early academic achievement. *Early Childhood Research Quarterly*, *28*(4), 774-783.

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, R., Augustyn, A., Klin, A., & Volkmar, F. (2005). Perception and production of prosody by speakers with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *35*, 205–220.

Peppé, S., & McCann, J. (2003). Assessing intonation and prosody in children with atypical language development: the PEPS- C test and the revised version. *Clinical Linguistics & Phonetics*, *17*(4-5), 345-354.

Peppé, S., McCann, J., Gibbon, J., O'Hare, A., & Rutherford, M. (2007). Receptive and expressive prosodic ability in children with high-functioning autism. *Journal of Speech, Language, and Hearing Research*, *50*, 1015–1028.

R Core Team (2022). R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Rapin, I., & Dunn, M. (2003). Update on the language disorders of individuals on the autistic spectrum. *Brain and Development*, *25*(3), 166-172.

Revelle, W. (2019). *psych: Procedures for Psychological, Psychometric, and Personality Research*. Northwestern University, Evanston, Illinois. R package version 1.9.12, https://CRAN.R-project.org/package=psych.

Roberts, J. A, Rice, M. L., & Tager-Flusberg, H. (2004). Tense marking in children with autism. *Applied Psycholinguistics*, *25*, 429–448.

Su, Y., & Naigles, L. R. (2019). Online Processing of Subject–Verb–Object Order in a Diverse Sample of Mandarin- Exposed Preschool Children with Autism Spectrum Disorder. *Autism Research*, *12*(12), 1829-1844.

Su, Y., & Naigles, L. R. (2021). Comprehension of grammatical aspect markers le and zai in a diverse sample of mandarin-exposed preschool children with autism spectrum disorder. *Reading and Writing*, 1-24.

Szendrői, K. (2004). Acquisition evidence for an interface theory of focus. *LOT Occasional Series*, *3*, 457-468.

Szendrői, K., Bernard, C., Berger, F., Gervain, J., & Höhle, B. (2018). Acquisition of prosodic focus marking by English, French, and German three-, four-, five-and six-year-olds. *Journal of Child Language*, *45*(1), 219-241.

Tager-Flusberg, H., Paul, R., & Lord, C. (2005). Language and communication in autism. In F. R. Volkmar, R. Paul, A. Klin, & D. Cohen (Eds.), *Handbook of autism and pervasive developmental disorders* (3rd ed., pp. 335–364). New York: Wiley.

Terzi, A., Marinis, T., & Francis, K. (2016). The interface of syntax with pragmatics and prosody in children with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *46*(8), 2692-2706.

Tovar, A. T., Fein D., & Naigles L. R. (2015). Grammatical aspect is a strength in the language comprehension of young children with autism spectrum disorder. *Journal of Speech, Language, and Hearing Research*, *58*(2), 301–310.

Wang, L., Beaman, C. P., Jiang, C., & Liu, F. (2022). Perception and production of statement-question intonation in autism spectrum disorder: A developmental investigation. Journal of Autism and Developmental Disorders, *52*(8), 3456-3472.

Wu, W., & Xu, Y. (2010). Prosodic focus in Hong Kong Cantonese without post-focus compression. *Proceedings of the 5th International Conference Speech Prosody*. Chicago, 1-4.

Zhou, F., Chen, S., Chan, A., Chun, E., Li, B., Tang, T., Choi, P. (2021) Mapping prosody and meaning by Cantonese-speaking children with and without autism spectrum disorder. *Proc. 1st International Conference on Tone and Intonation (TAI)*, 186-189, doi: 10.21437/TAI.2021-38.

Zhou, P., & Crain, S. (2010). Focus identification in child Mandarin. *Journal of Child Language*, *37*(5), 965-1005.

Zhou, P., Crain, S., Gao, L., Tang, Y., & Jia, M. (2015). The use of grammatical morphemes by Mandarin-speaking children with high functioning autism. *Journal of Autism and Developmental Disorders*, *45*(5), 1428-1436.

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

Zimmermann, M., & Onea, E. (2011). Focus marking and focus interpretation. *Lingua*, *121*(11), 1651-1670.

Figure Captions

Figure 1. Example of the picture stimuli for the experiment

Figure 2. Mean percentage of YES response in Cantonese-speaking adults, children with ASD and TD children in experiment. Error bars indicate ±1SE.

Figure 3. Mean log-transformed RTs in Cantonese-speaking adults, children with ASD and TD children in experiment. Error bars indicate ± 1SE.

Figures

Figure 1 top. Example of the picture stimuli for the experiment

Figure 2 top. Mean percentage of YES response in Cantonese-speaking adults, children with ASD and TD children in experiment. Error bars indicate ±1SE.

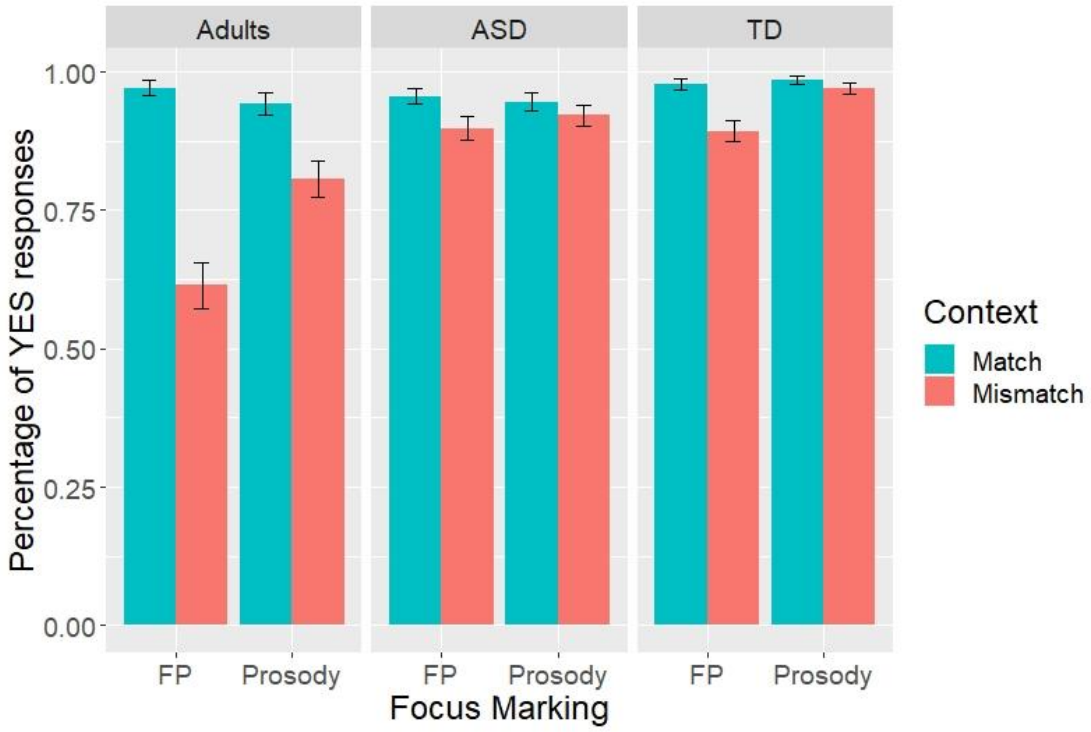
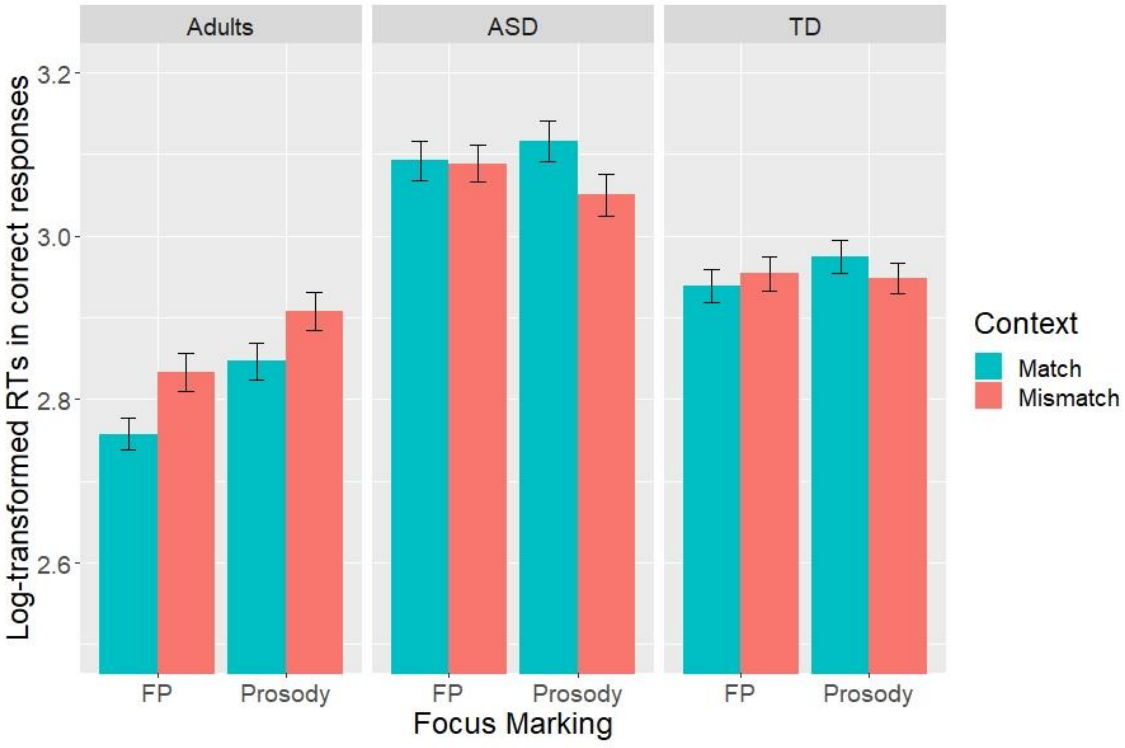


Figure 3 top. Mean log-transformed RTs in Cantonese-speaking adults, children with ASD and TD children in experiment. Error bars indicate ± 1SE.



Tables

Table 1. Information of participants (SD in parentheses) and results of Primary Test of Nonverbal Intelligence (PTONI) test, Cantonese Receptive Vocabulary Test (CRVT) and working memory (WM) by Cantonese-speaking children with ASD and TD children.

						PTONI				CRVT [a]		WM	
		Age				Raw scores		Nonverbal index		Raw scores		Raw scores	
Group	Number	*Mean*	*Range*	Female	Male	*Mean*	Range	*Mean*	Range	*Mean*	Range	*Mean*	Range
ASD	44	7;00 (1.17)	5;00 – 8;10	8	36	41.295 (12.315)	19 – 67	111.591 (24.069)	59 – 143	59.318 (5.822)	43 – 65	5.300 (2.053)	2 – 10
TD	54	6;03 (1.09)	5;00 – 8;11	20	34	43.442 (11.459)	23 – 66	114.330 (21.573)	71 – 149	61.173 (2.826)	53 – 65	5.191 (2.163)	2 – 10
Adults	28	19.61 (1.91)	18;00 – 25;00	22	6								

Note. [a] CRVT does not provide standard scores.

Table 2. Example of experimental stimuli

Question:						
邊個	食緊	條	蘿蔔	啊？		
Who	eat-PROG	CL	carrot	SFP		
"Who is eating the carrot?"						
我	覺得	馬騮	食緊	條	蘿蔔。	
I	think	monkey	eat-PROG	CL	carrot	
"I think the monkey is eating the carrot."						
Answer:						
Condition (a)	唔係，	係	[狐狸]$_F$	食緊	條	蘿蔔。
	No,	FP	fox	eat-PROG	CL	carrot
Condition (b)	唔係，	[狐狸]$_F$	食緊	條	蘿蔔。	
	No,	fox	eat-PROG	CL	carrot	
Condition (c)	唔係，	狐狸	食緊	係	條	[蘿蔔]$_F$。
	No,	fox	eat-PROG	FP	CL	carrot
Condition (d)	唔係，	狐狸	食緊	條	[蘿蔔]$_F$	
	No,	fox	eat-PROG	CL	carrot	
	"No, the fox is eating the carrot."					

Table 3. ANOVA table of the final model for YES–NO judgments of Cantonese-speaking adults, children with ASD and TD children

Model: Group*FocusMarking*Context + (1|Item)+(1|Participant)

	LR Chisq	***Df***	***p***
Group	9.992	2	**<0.001**
Focus Marking	20.269	1	**<0.001**
Context	132.709	1	**<0.001**
Group : Focus Marking	4.290	2	**0.014**
Group : Context	39.650	2	**<0.001**
Focus Marking : Context	30.873	1	**<0.001**
Group : Focus Marking : Context	7.214	2	**<0.001**

Table 4. Mean RTs (ms) (SD in parentheses) of Cantonese-speaking adults, children with ASD and TD children in four conditions.

	Adults	**Children with ASD**	**TD children**
Condition (a)	688.04 (465.99)	1709.29 (1441.56)	1204.77 (1066.29)
Condition (b)	895.02 (691.11)	1800.46 (1557.43)	1288.02 (1098.21)
Condition (c)	864.69 (638.39)	1612.55 (1263.49)	1348.21 (1227.15)
Condition (d)	998.65 (677.70)	1598.81 (1469.27)	1231.65 (1140.31)

Note: Condition (a) focus marked by FP in matched context; Condition (b) focus marked by prosody with matched context; Condition (c) focus marked by FP in mismatched context; Condition (d) focus marked by prosody in mismatched context

Table 5. ANOVA table of the final model for log-transferred RTs in correct response of Cantonese-speaking adults, children with ASD and TD children

Model: Group + FocusMarking + Context + Group*FocusMarking + Group*Context + FocusMarking*Context + (1|Item)+(1|Participant)

	LR Chisq	***Df***	***p***
Group	20.536	2	**<0.001**
Focus Marking	1.692	1	0.193
Context	1.979	1	0.159
Group : Focus Marking	2.405	2	0.091
Group : Context	39.650	2	**0.004**
Focus Marking : Context	30.873	1	**0.039**

Author Note

Haoyan GE, School of Education and Languages, Hong Kong Metropolitan University, Hong Kong SAR, China

Fang LIU, School of Psychology & Clinical Language Sciences, The University of Reading, UK

Hoi Kwan YUEN, School of Education and Languages, Hong Kong Metropolitan University, Hong Kong SAR, China

Aishu CHEN, School of English Education, Guangdong University of Foreign Studies, China

Virginia YIP, Department of Linguistics and Modern Languages & Childhood Bilingualism Research Centre, The Chinese University of Hong Kong, Hong Kong SAR, China

The work described in this paper was substantially supported by two grants to the first author from the Hong Kong Research Grants Council – Faculty Development Scheme (Project Reference No.: UGC/FDS16/H13/19), and Katie Shu Sui Pui Charitable Trust – Research and Publication Fund (Project Reference No.: KS2018/2.2). FL was supported by a European Research Council (ERC) Starting Grant (CAASD, 678733).

Correspondence concerning this article should be addressed to Haoyan Ge, School of Education and Languages, Hong Kong Metropolitan University, Ho Man Tin, Kowloon, Hong Kong SAR, China. E-mail: hge@hkmu.edu.hk