

Coconuts and curtain cakes: The production of *wh*-questions in ASD

Nufar Sukenik, Eleonor Morin, Naama Friedmann, Philippe Prevost, & Laurice Tuller

Abstract

Background and aims. Children with autism spectrum disorders (ASD) have been found to exhibit difficulties in *wh*-question production. It is unclear whether these difficulties are pragmatic or syntactic in nature. The current study used a question elicitation task to assess the production of subject and object *wh*-questions of children with ASD in two different languages (Hebrew and French) wherein the syntactic structure of *wh*-questions is different, a fact that may contribute to better understanding of the underlying deficits affecting *wh*-question production.

Crucially, beyond the general correct / error rate we also performed an in-depth analysis of error types, comparing syntactic to pragmatic errors and comparing the distribution of errors in the ASD group to that of children with typical development (TD) and children with Developmental Language Disorder (DLD).

Results. Correct production rates were found to be similar for the ASD and DLD groups, but error analysis revealed important differences between the ASD groups in the two languages and the DLD group. Both ASD groups were found to produce similar pragmatic errors, which were not found in children with DLD. Syntactic errors were affected by the structure of each language.

Conclusions. Our results have shown that although the two ASD groups come from different countries and speak different languages, the correct production rates and more importantly, the error types were very similar in the two ASD groups, and very different compared to TD children and children with DLD.

Implications. Our results highlight the importance of creating research tasks that test different linguistic functions independently and strengthen the need for conducting fine-grained error analysis to differentiate between groups and gain insights into the deficits underlying each of them.

Introduction

“Coconut”. This was the response of JOC, one of our French-speaking participants with Autism Spectrum Disorder (ASD) when we prompted him to form the question “Who are you pushing?”. “What is behind the curtain cake?” was yet another unusual response, from a Hebrew-speaking participant with ASD. Children with ASD have been reported to perform below their age-matched typically developing (TD) peers in the comprehension and production of *wh*-questions. There is ongoing debate regarding the nature of the deficits that cause *wh*-questions to be so elusive for children with ASD. Some consider it to be a result of a syntactic impairment, which affects syntactically complex structures, while others attribute these difficulties to impaired pragmatics. In the current study we assessed the ability of Hebrew and French-speaking school-aged children and adolescents with ASD to produce *wh*-questions on an elicited production task and explored in detail their response patterns, with the aim of investigating the nature of their difficulties.

Wh-questions in Hebrew and in French

Wh-questions contain a *wh*-phrase (*who*, *which girl*, *why*, etc.), which can correspond to a subject, an object, or an adjunct. *Wh*-questions in languages such as English or Hebrew are obligatorily derived by *wh*-movement in which the *wh*-phrase moves to the beginning of the clause; in other languages, such *wh*-movement never takes place (e.g., Korean), and in some it is only optional (e.g., French).

The current study focuses on *wh*-questions in French and Hebrew, languages in which the syntax of *wh*-questions is quite different. Hebrew *wh*-questions involve obligatory *wh*-movement of the *wh*-phrase to a clause-initial position which may be accompanied by movement of the verb (so-called subject-verb inversion or triggered inversion, Shlonsky, 1997). Reznick and Friedmann (2017) found that TD children as young as a year and half already produce various kinds of *wh*-questions with syntactic movement – subject and object *who* questions, *where* questions, although at the very beginning these are mainly phrases that were learned as a whole (*ma-ze* ‘what-is-that?’, Armon-Lotem, 2008). Some types of structures, and questions in particular, are more difficult than others in cases of language impairment, and are acquired later than the others in typical acquisition. One such factor that crucially affects acquisition and impairment is whether or not the structure involves an object that moves to the beginning of the sentence across the subject, and specifically when the subject and object are full noun phrases (such as "the elephant" or "the girl"). (More formally, such structures are intervention configurations in which a lexically-restricted DP crosses another lexically-restricted DP, Friedmann, Belletti, & Rizzi, 2009). For example, in an object question like "which girl does the grandma draw", the object, "which girl" moves across the subject "the grandma", creating an object question with intervention); in the parallel subject question "which grandma draws the girl", no such crossing movement occurs). Reznick and Friedmann (2017) found that until age 5-6 years, the questions that children acquiring Hebrew produce do not involve such intervention configurations, neither can they understand such questions at this age. This means that they have difficulty in producing and understanding object questions with two full noun phrases, which involve such intervention configuration, but no problem in subject questions, which do not involve intervention (Friedmann, Belletti, & Rizzi, 2009; Biran & Ruigendijk, 2015). Around age 6-7 Hebrew-speaking children already master the comprehension and production of *wh*-questions, including questions that involve such intervention configurations such as object questions with two full noun phrases (Friedmann & Szterman, 2011; Friedmann & Novogrodsky, 2011).

In French, *wh*-questions can be formed in several different syntactic ways, varying in terms of the number of movement operations involved, from no *wh*-related movement at all, in the case of *wh*-in situ questions, to *wh*-movement accompanied by V-movement, and including, in between, *wh*-movement with no V-movement (as is possible in Hebrew) and *wh*-movement with a question morpheme immediately following the moved *wh*-phrase (see Prévost et al., 2010). Various groups of children acquiring French, including young TD children, children with a developmental disorder affecting language acquisition, such as Developmental Language Disorder, as well as children acquiring French as a second language, produce high rates of *wh* in situ *wh*-questions compared to *wh*-questions involving movement (Hamann 2006; Jakubowicz, 2011; Prévost et al., 2010). These results have been interpreted as evidence for avoidance of movement operations.

Wh-questions in children with ASD

Children with autism have been found to exhibit deficits in several complex syntactic constructions such as passives (Durrleman et al., 2017; Gavarro & Heshmati, 2014; Terzi et al., 2014), relative clauses (Durrleman et al., 2016; Sukenik & Friedmann, 2018) and pronominal clitics (Durrleman & Delage, 2016; Terzi et al., 2014; Tuller et al., 2017). *Wh*-questions, as described in the previous section, are also complex syntactic structures. Studies reporting that *wh*-questions are hard for children with ASD generally explained this difficulty in terms of the computational complexity involved in the syntactic operations underlying their derivation (Durrleman et al., 2016; Prévost et al., 2018; Tuller et al., 2017).

Several recent studies testing *wh*-question comprehension in children with ASD found their performance to be lower than that of age-matched TD peers (Durrleman et al., 2016; Jyotishi et al., 2017; Su et al., 2014). Moreover, while in some studies ASD performance seemed to improve with age, in others this was not the case. Durrleman et al. (2016) found that syntactic complexity seems to affect how children with ASD comprehend *wh*-questions (and relative clauses) in the same way as TD children. Their study moreover found that in TD children there seemed to be improvement with age, whereas in the ASD group performance on syntactically complex structures was found to be related to their nonverbal IQ (NVIQ). Jyotishi et al. (2017) found that the comprehension of object *wh*-questions emerged later in children with ASD compared to their TD peers, but at similar levels of overall language scores. Su et al. (2014) found an age effect, with older children with ASD achieving higher scores on *wh*-question comprehension.

Several studies have also noted that some children with ASD perform oddly on tasks designed to assess complex syntactic structures, such as those involving syntactic movement, in that their errors appear not to be only syntactic in nature. These types of errors have not been reported for children with syntactic Developmental Language Disorder (DLD); they seem to stem from a social/pragmatic deficit (Demouy et al., 2011; McGregor et al., 2012; Modyanova et al., 2017; Naigles et al., 2016; Sukenik & Friedmann, 2018; Roberts et al., 2004; Williams et al., 2008). Some studies have suggested that these odd, pragmatically deviant responses of children with ASD could be a marker distinguishing ASD from DLD (Roberts et al., 2004; Williams et al., 2008).

One clue as to whether the difficulty with *wh*-questions results from a syntactic or a pragmatic deficit comes from Korean. A good candidate for the basis of a syntactic deficit in *wh*-questions is the movement of the *wh*-element. In Korean, *wh*-constituents remain in situ, so Korean *wh*-questions do not involve *wh*-movement, thus making them less complex than *wh*-questions in English, which involve movement. Therefore, if the ASD children do not show difficulties with *wh*-questions in Korean, this would suggest that the deficit is related to syntactic movement. If, however, they still show difficulties even though the *wh*-questions they need to understand do not involve movement, this would lend support for a pragmatic source of the difficulty. Park (2016) tested *wh*-question comprehension in Korean-speaking children with ASD. The findings were that the ASD children nevertheless had poorer comprehension of *wh*-questions than TD age-matched controls. These findings, according to which ASD children have difficulties even in Korean *wh*-questions, which do not involve syntactic movement (as the *wh*-constituent remains in situ), support the conclusion that pragmatics may (also) a source of difficulty.

Goodwin et al. (2012) addressed the issue of a potential syntactic-pragmatic confound in interpreting performance on *wh*-questions in children with ASD. They argued that teasing these two apart is possible, for comprehension, through the use of an intermodal preferential looking task. They compared the emergence of various types of *wh*-questions in comprehension of children with ASD to their emergence in production, which was analysed in semi-structured play speech samples. They found that just like the TD children, the children with ASD displayed comprehension of a given type of *wh*-question before they began to produce such questions and that the comprehension of questions appeared at the same linguistic levels as it did in TD control children, albeit at an older chronological age. Although comprehension levels were found to be strongly linked to overall linguistic levels, the majority of ASD participants in this study still failed to produce *wh*-questions in free speech. The authors listed several reasons that children with ASD may produce fewer *wh*-

questions than their typically developing peers: (1) problems with complex linguistic structures (a syntactic reason), (2) problems understanding that the *wh*-word stands for information that isn't present in the spoken sentence (abstract reasoning), (3) communication deficits that results in not knowing when to ask questions appropriately (pragmatic reason), and (4) not understanding someone else's needs to answer the question (pragmatic or theory of mind related reason).

Another explanation for the low levels of *wh*-question production may be the interaction between the acquisition of syntax and pragmatics. In typical development the different language domains impact each other and develop in synchrony (Hoff, 2013). The language of children with ASD, on the other hand, is known to be asynchronic at times (Eigsti & Schuh, 2017). Therefore, it could be that a syntactic deficit (that seems common in ASD individuals) may affect the acquisition of pragmatic rules, leading to reduced *wh*-question production. The other way around may also occur, where the well documented pragmatic deficits of children with ASD affect their syntactic acquisition again leading to lower levels of *wh*-question production.

Summarizing, while it seems that many individuals with ASD have trouble producing complex syntactic structures such as *wh*-questions, it is unclear whether the underlying deficits associated with *wh*-question production are syntactic or pragmatic in nature.

Research questions

The current study stems from two independent studies of elicited *wh*-questions in children with ASD, one conducted on Hebrew-speaking children in Tel Aviv, and the other conducted on French-speaking children in Tours. The same basic *wh*-question elicitation task had been administered in both labs as part of wider investigations of language skills of children with ASD (Prévost et al., 2018; Sukenik, 2012; Sukenik & Friedmann, 2018; Tuller et al., 2017; Zebib et al., 2013). The motivation for the current comparative study was twofold. First, the syntax of *wh*-questions in French and Hebrew is quite different. As reviewed above, French-speaking children with ASD have at their disposal a very simple syntactic strategy (*wh*-in situ), which is not available to Hebrew-speaking children. All other things being equal, we might expect that, if the major difficulty in producing *wh*-questions is grammatical in nature, then the French-speaking children might have fewer difficulties producing *wh*-questions than the Hebrew-speaking children because of the availability of *wh*-in situ in French, but not in Hebrew. They may also produce more in-situ questions than questions that involve movement of the *wh*-phrase, and produce more questions without intervention (such as subject questions) than questions with intervention (such as object questions, in which the object moves

across the subject). The second reason for bringing together and subjecting to further analysis the *wh*-question productions of the children in the two studies was our discovery that both groups of children had produced a considerable number of odd responses which were strikingly similar to each other. The above-cited “Coconut!” and “What's behind the curtain cake?” responses were emblematic of these types of responses.

Pooling our data and subjecting it to both detailed syntactic and pragmatic coding, we reasoned, could further our understanding of the nature of reported difficulties with question formation in children with ASD. Specifically, we sought to provide answers to the following questions: 1) How frequently do children with ASD produce expected *wh*-questions, ones that are syntactically and pragmatically well-formed? and 2) What is the nature of unexpected *wh*-questions produced by these children—are they due to syntactic or pragmatic deficits?

Method

We report on two independent question elicitation studies, one in French, the other in Hebrew, which were administered before we had the idea of comparing the findings. Hence, although the methods were similar and included the same pictures, they are not identical, and the characteristics of the children in the different language groups were different.

Participants

French-speaking participants

The group of French-speaking children with ASD (ASD-FR) consisted of 16 boys and 4 girls, all native monolingual French speakers, aged 6;3 to 12;9 years ($M = 8;7$ years, $SD = 1;9$ years). They had been diagnosed according to International Classification of Diseases (10th revision) criteria and confirmed via the Autism Diagnostic Interview Revised (ADI-R) (Lord et al., 1994) and by the Autism Diagnostic Observation Schedule (ADOS), module 2 or 3 (Lord et al., 1989). To ensure that language tests could be administered, only children producing sentences of at least two or three words, according to clinicians (speech language pathologists and/or doctors), were included. Morphosyntactic expressive skills were assessed for all of them via a standardized test (the sentence completion task from Khomsi et al., 2007) as well as, for 13 of them,¹ by measuring Mean Length of Utterances (MLU) in a spontaneous speech sample ($M = 5.81$, $SD = 1.38$). Many of these children displayed low expressive

¹ Three children were too tired or simply refused to talk with the experimenter.

morphosyntax scores, and most of them (13/16) attended speech-language therapy. Regarding cognitive abilities, no inclusionary criterion was set, voluntarily, for IQ, and this measure indeed varied greatly within the group (full-scale IQ² from 42 to 108). Raven's Progressive Matrices (RPM, Raven et al., 1986) were administered to determine nonverbal reasoning levels. Performance of ASD-FR children varied from severe nonverbal disability to normal nonverbal level ($M = 31.75$, $SD = 28.19$); see Prévost et al., (2017) and Tuller et al., (2018). As is typical for ASD (Lombardo et al., 2019), individual variation within the group of children with ASD was very large on all variables. Group scores do therefore not provide a clear picture of participant characteristics; Table 1 presents individual scores of the children in the ASD -FR group.

Table 1. French-speaking children with ASD: age, sex, nonverbal level and expressive morphosyntactic skills.

<i>Child</i>	<i>Age (months)</i>	<i>Sex</i>	<i>RPM percentile</i>	<i>Sentence Completion Z-score</i>	<i>MLU</i>
FRA	75	M	62.5	-1.0	6.68
JOC	76	M	2.5	-1.3	4.3
BRL	83	M	2.5	-2.0	3.85
WAE	83	M	75.0	-1.7	-
BRR	86	M	2.5	-1.9	4.71
LIK	87	F	50.0	0.1	4.54
HEG	89	F	92.5	-1.4	5.84
JUF	96	M	50.0	-2.2	6.95
LUJ	97	M	7.5	-1.2	5.92
MAV	97	M	37.5	-1.2	5.86
ARE	103	M	7.5	-1.8	6.00
ETG	105	M	17.5	-3.1	5.47
AUJ	110	M	37.5	-3.1	-
MAD	110	M	37.5	-2.0	4.89
SEG	112	F	5.0	-2.7	-
MAM	114	F	10.0	-4.7	5.28
ROD	122	M	7.5	-3.4	6.52
KIH	125	M	75.0	-0.8	8.97
HEJ	137	M	17.5	0.8	8.33
ARF	153	M	37.5	-5.3	4.71

² Children were tested using the Wechsler Intelligence Scale for Children (WISC), Wechsler Preschool and Primary Scale of Intelligence, Kaufman Assessment Battery for Children or Differential Scales of Intellectual Efficiency.

In order to provide a comparative perspective, these French children with ASD were age-matched to 20 children with phonological-syntactic Developmental Language Disorder (DLD-FR) (12 boys and 8 girls, aged 6;5 to 10;7 years, $M = 8;7$ years, $SD = 1;5$ years). The children with DLD had all received a conventional diagnosis based on severely impaired language performance on standardized language batteries commonly used by speech and language therapists in France, in conjunction with usual exclusionary criteria, including intellectual disability (as measured by the Wechsler Intelligence Scale for Children [performance IQ < 80], or an equivalent evaluation) and autism (routine diagnostic procedure included psychiatric examination). These children had minimum MLUs of 2.5. The ASD group and the DLD group did not differ in age ($U(40) = 209, p = .82$).

French children with ASD were also compared to a group of 12 six-year-old typically developing children (TD-FR) ($M = 6;5, SD = .2$). The TD children were thus the same age as the youngest children in the ASD group; our reasoning was that only TD children this young could be argued to be susceptible to producing pragmatically inappropriate responses, and we furthermore expected that the morphosyntactic abilities needed to produce *wh*-questions are fully acquired at this age.

All of the French-speaking participants were recruited as part of a larger research project supported by the French National Research Agency (Grant ANR BLAN_0328-01) in close collaboration with speech-language pathologists and doctors at autism centers in university teaching hospitals in Tours and in Brest and at the Language Reference Center of the University Children's hospital in Tours. TD participants were recruited and tested in local public schools. University ethics procedures for research involving human subjects were strictly adhered to.

Hebrew-speaking participants

The Hebrew-speaking participants with ASD (ASD-HE) consisted of 18 native Hebrew speakers.³ They were aged 9;0 to 18;0 years ($M = 13;4$ years, $SD = 3;1$ years), 16 boys and 2 girls, who were diagnosed with autism prior to the study, according to the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV, American Psychiatric Association, 2000) criteria, by a psychiatrist. Thirteen were diagnosed with high-functioning autism (HFA), and five were diagnosed with ASD. None had a diagnosis of intellectual disability or other learning disability. All children in this group were enrolled in special education classes for children with ASD integrated in regular schools across Israel. All children who

³ One child was a native speaker of both Hebrew and English.

agreed to participate in the study were included. Nonverbal intelligence was assessed by a picture-matching association task developed at the Language and Brain Lab at Tel Aviv University (MA KASHUR picture association test, Biran & Friedmann, 2007). The aim of the task was to test the conceptual system and the way the participants understood relations between objects. Morphosyntactic measures included a sentence repetition task targeting syntactically complex structures (PETEL, Friedmann, 2000) developed at the Brain and Language Lab at Tel Aviv University. The task consisted of 70 sentences that the participant was asked to listen to and repeat. Table 2 displays the ASD children's scores, with boldfaced scores indicating scores that were significantly lower than age controls. It can be observed that nonverbal scores of the ASD-HE participants were generally high, but many had low scores on the sentence repetition task regardless of their age.

Table 2. Hebrew-Speaking children with ASD, age, sex, diagnosis, Nonverbal task and expressive morphosyntactic.

<i>Child</i>	<i>Age (months)</i>	<i>Sex</i>	<i>NV- association task (% correct)</i>	<i>Sentence repetition (% correct)</i>
OG	115	M	92*	76
AL	121	M	100	70
AO	123	M	95	84
TZ	127	M	95	20
TR	136	F	100	86
LA	140	M	97	54
AC	142	M	95	86
RH	144	F	100	39
TW	147	M	89	81
YO	190	M	100	91
AA	192	M	100	41
AS	193	M	100	91
MS	194	M	100	6
EE	196	M	97	84
ME	198	M	100	71
NA	204	M	100	70
NZ	219	M	100	96
AG	222	M	100	91

Notes: *Bold-faced scores are scores that were significantly lower than TD controls.

** HFA- children diagnosed prior to the current study with High Functioning Autism, typically defined as IQ scores > 75.

The ASD children were compared to a group of 12 typically-developing children (TD-HE) with no known disabilities aged 7;5-9;0 ($M = 8;3$ years, $SD = 0.6$), most of whom were younger than the ASD participants. Comprehension and production of all types of *wh*-questions in Hebrew is fully acquired around age six (Friedmann & Novogrodsky, 2011; Friedmann et al., 2009; Friedmann & Szterman, 2011; Reznick & Friedmann, 2017). We therefore reasoned that children aged seven years would most likely have mastered comprehension and production of *wh*-questions, but may still be novice enough to produce various errors that would not be observed in more experienced speakers.

This study was part of a larger study assessing language abilities of children with ASD at the Brain and Language lab at Tel Aviv University. Ethics approval was attained from the Tel Aviv University Ethics Committee and the Ministry of Education. School principals were contacted and asked to participate in the study. Out of 20 schools contacted, three agreed to participate. Next, recruitment letters were sent out to parents via the teachers (return rate was about 80%). Tasks were all administered by a graduate student who was a certified special education teacher. All testing sessions were conducted in the morning hours at the child's school.

Materials and testing procedure

The Hebrew and the French studies used a similar task format that is commonly used in psycholinguistic research to test question formation; it is also part of at least one standardized language battery (Seymour et al., 2003), as a test of pragmatic abilities. Both tasks elicited subject and object *wh*-questions on the basis of a set of common pictures (taken from BAFLA, Friedmann, 1998). Examples of these pictures are provided in Figures 1 and 2. Each picture showed an action being performed by one participant on another. The action (a verb in the singular present form, which would be clearly depicted in a picture, such as washing, biting, kissing) was always visible, but either the agent or the theme was concealed by an opaque silhouette (white in the French task and silver in the Hebrew one). The children were required to ask a question for which the answer would be the concealed figure. Right after the child's answer (*wh*-question or other response), the experimenter showed an identical picture which revealed the hidden figure.



Figure 1: A picture used to elicit a subject question: “Who is biting you?” (French test)

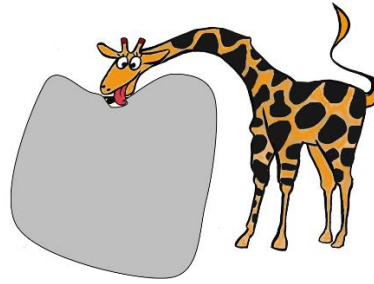


Figure 2: A picture used to elicit an animate object question: “Who is the giraffe licking?” (Hebrew test)

Both the French and the Hebrew versions of the experiment elicited production of subject and object *who* questions. The French version also elicited *what* object questions and adjunct *wh*-questions (*how*, *why*, and *where*). We report here only on subject and object questions: in the Hebrew test, there were 20 subject *who* questions and 20 object *who* questions, and, in the French test, there were 12 subject questions, 12 animate object (*who*) questions, and 12 inanimate object (*what*) questions.

The procedures in the two experiments were somewhat different. In Hebrew, the experiment started with an example picture for which the children were asked to produce a question and the experimenter explained the task. The experimenter emphasized that the concealed figure is always animate – it is either a person or an animal (both triggering the same animate *wh*-morpheme, *mi*, 'who' in Hebrew). During the test itself there were no lead-in sentences for each picture, and thus neither the *wh*-word nor the verb was provided to the child. (1) and (2) are the expected Hebrew questions for Figures 1 and 2, respectively.

(1) Hebrew subject *who* question (Figure 1)

Expected response: mi noshet et ha-kelev?
 who bites ACC the-dog
 ‘Who is biting the dog?’

(2) Hebrew object *who* question (Figure 2)

Expected response: et mi ha-arnav doxef?
 ACC who the-rabbit pushes
 ‘Whom is the rabbit pushing?’

In the French version, each picture was accompanied by a sentence about the visible participant in the picture, and the child was led to ask this participant a question, as illustrated in (3) and (4), the French items accompanying Figures 1 and 2, respectively.

(3) Experimenter: Quelqu'un mord le chien, mais on ne voit pas qui. Demande-lui.

'Someone is biting the dog, but we can't see who. Ask him!'

Expected response: Qui te mord?

who you bites

'Who is biting you?'

(4) Experimenter: Le lapin pousse quelqu'un, mais on ne voit pas qui. Demande-lui.

'The rabbit is pushing someone, but we can't see who. Ask him!'

Expected response: Tu pousses qui? / Qui tu pousses? /

you push who? / who you push?

Qui est-ce que tu pousses? / Qui pousses-tu?

who ESK (=question marker) you push who push you

'Who are you pushing?'

As can be seen in these examples, the two tasks also differed in the way in which the child addresses the character: in the Hebrew task, the child didn't have to address the character directly whereas in the French task he was supposed to do so. Consequently, in the French task, the child had to use a pronoun (subject or object clitic).

In both studies, participants were tested individually in a quiet room. The children participated voluntarily in the experiments and they were told that they could stop whenever they wanted. No time limit was imposed. The experimenter gave general encouragement. If the participant looked tired or asked for a break the test was stopped and continued at the next meeting. All sessions were audio-recorded for later transcription and coding.

Data Analysis

In order to determine the prevalence and the nature of children's difficulties producing *wh*-questions, we developed a step-by-step coding procedure, outlined in Figure 3.

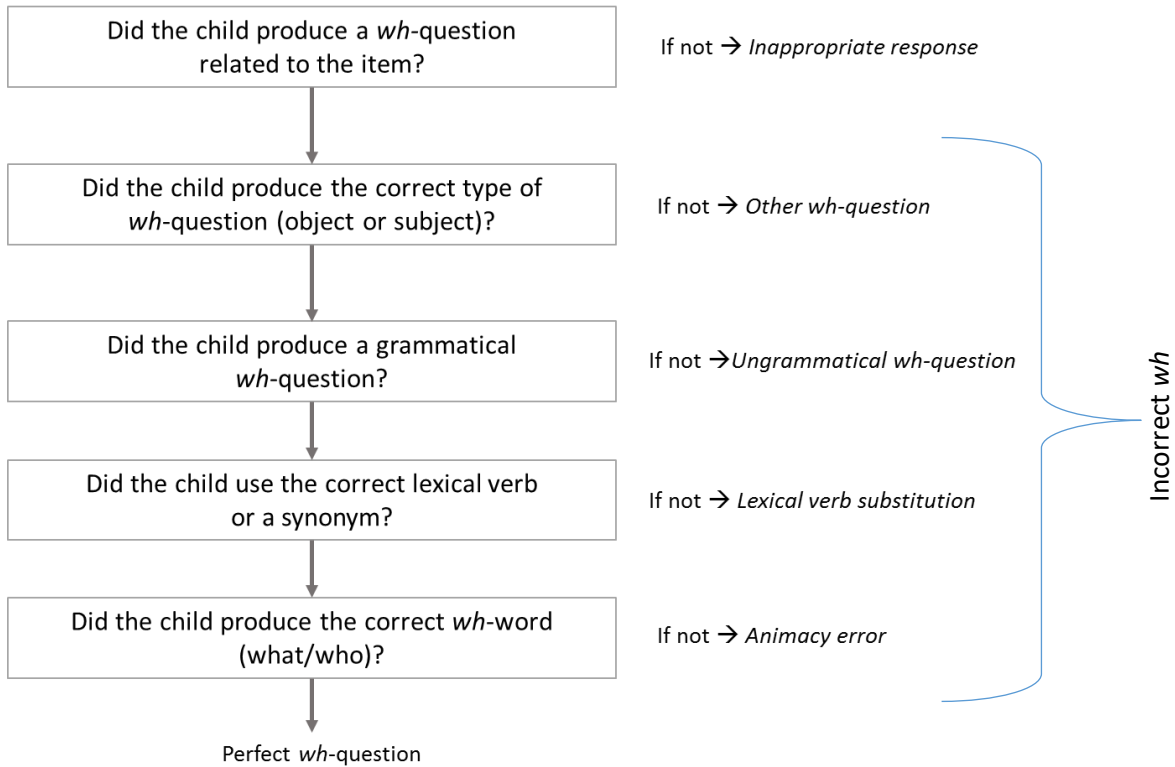


Figure 3: Step-by-step coding procedure

Table 3: Examples of responses, coding scheme and labels, corresponding to Figure 2, for the animate object question “Who are you pushing?”

<i>Child response</i>	<i>Appropriate wh-question</i>	<i>Correct type of wh-question</i>	<i>Grammatical wh-question</i>	<i>Correct lexical verb</i>	<i>Correct animacy</i>	<i>Label</i>
Est-ce que poussez? 'Are you pushing?'	0	-	-	-	-	Inappropriate response
Qu'est-ce que tu fais? 'What are you doing?'	0	-	-	-	-	Inappropriate response
C'est peut-être un sac! 'It could be a bag!'	0	-	-	-	-	Inappropriate response
Qui te pousse? 'Who is pushing you?'	1	0	-	-	-	Other <i>wh</i> -question
Qui toi pousser? 'Who you push?'	1	1	0	-	-	Ungrammatical <i>wh</i> -question
Qui tu embrasses? 'Who are you kissing?'	1	1	1	0	-	Lexical verb substitution
Qu'est-ce que tu poussez? 'What are you pushing?'	1	1	1	1	0	Animacy error
Qui tu poussez? 'Who are you pushing?'	1	1	1	1	1	Perfect <i>wh</i> -question

First, we determined whether the child produced a *wh*-question that was pragmatically appropriate to the item, mentioning both the correct action and the participants in that action. If the child did not produce such a *wh*-question, his/her production was coded as an inappropriate response and was excluded from all further syntactic analysis. For productions that consisted of a *wh*-question related to the item, we then checked whether the child produced the right kind of *wh*-question (object question or subject question), whether the *wh*-question was grammatical, whether the child used the correct lexical verb or a synonym, and whether the animacy of the *wh*-word was correct. This coding procedure led to different types of responses, illustrated in Table 3. The responses were either Perfect *wh*-questions, Inappropriate responses, Incorrect *wh*-questions, which included Ungrammatical *wh*-questions, Other *wh*-questions, Lexical verb substitutions, and Animacy errors. As these responses were similar in the two languages, we present an example of the different responses for Figure 2 in only one language, French.

In each language, the correct production rates and the different rates of the various types of errors were compared to the production rates of TD children, and in the French group also to DLD children.

Analytic strategy

In order to test the homogeneity of the task results we used Levene's test. Levene's test showed that the variances in response patterns between the groups were not equal $F(3,78) = 6.01, p = .001$. Next we used a Shapiro-Wilk to test for normality and found that in all groups there was a significant departure from normality, $W(82) = .91, p < .001$. Following these results, we used non-parametric tests for all statistical analyses. The Mann-Whitney U Test was used for comparing performance between ASD and TD groups in HE and FR, and between the ASD and DLD groups (FR). Because in the French study we performed two preplanned comparisons of the performance of the ASD group – with each of the control groups (TD, DLD), we used FDR correction (Benjamini & Hochberg, 1995). A correlation analysis based on Spearman's rank correlation coefficient (r_s) was used to test whether any of the background measures in the ASD groups (HE, FR) was related to the rates of correct *wh*-question production rates.

All data were transcribed and coded by at least two French-speaking expert linguists or two Hebrew-speaking expert linguists. All disagreements were discussed and resolved together, as was coding of the error types. Error coding, developed specifically for this study, was first done by each language team separately and then reviewed a third time by all authors together in order to resolve any unclear errors by consensus of all researchers in the team.

Results

General rate of perfect *wh*-questions

A first analysis was conducted by counting the production rates for perfect *wh*-questions (Figure 4 presents the correct percentages of each group, except for the TD-HE group who performed at ceiling levels). Perfect *wh*-questions were questions in which the participant produced the correct syntactic form, the correct *wh*-word, the correct verb (or a synonym) and mentioned the correct participants.

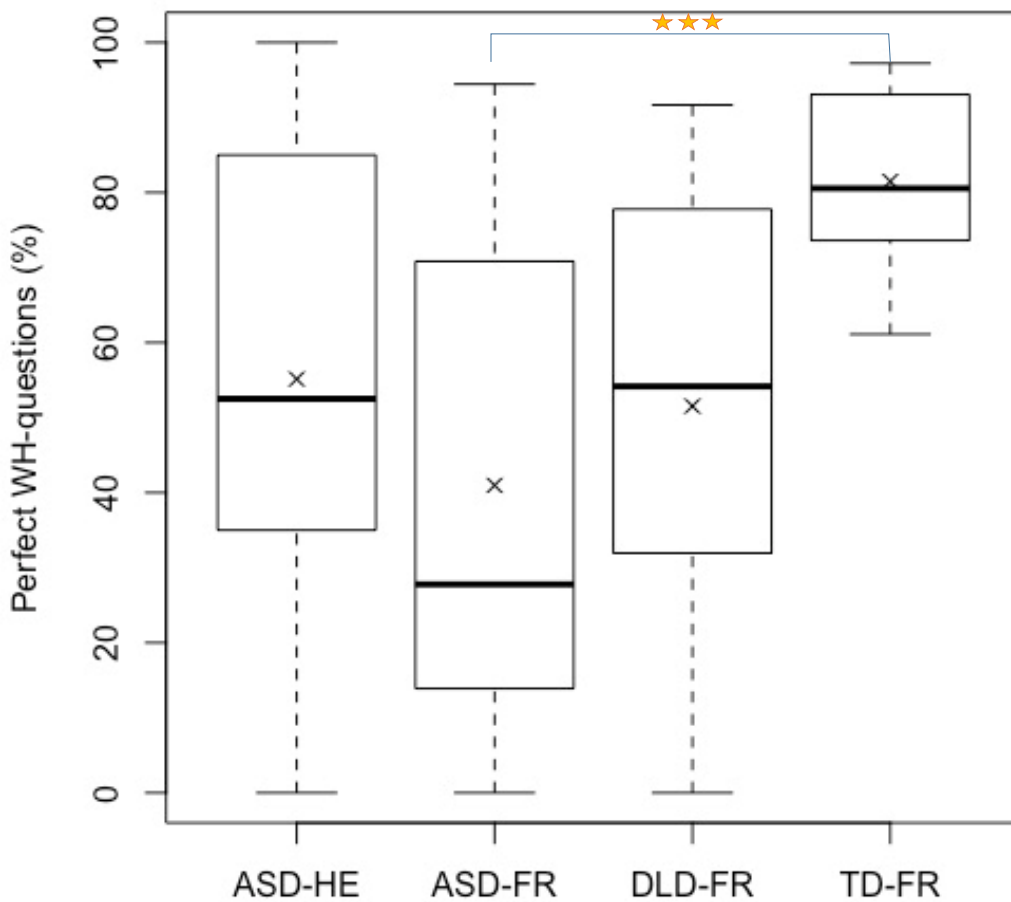


Figure 4: Production rates (% correct) for perfect *wh*-questions in the four participant groups, median (bar) and mean (cross). Box and whiskers indicate quartiles based on the mean. *** Indicate a significant difference of $p < .001$.

As laid out in Figure 4, there was more variability in the performance of the three clinical groups (ASD-HE, ASD-FR and DLD-FR) than in the TD groups (TD-FR and TD-HE). Furthermore, mean percentages of perfect *wh*-questions among the three clinical groups were similar (ASD-HE: $M = 55.1\%$, $SD = 26.1$; ASD-FR: $M = 41.0\%$, $SD = 29.7$; DLD-FR: $M = 51.5\%$, $SD = 28.8$) and much lower than those of the TD children (TD-FR: $M = 81.5\%$, $SD = 11.9$; TD-HE: $M = 99.6\%$, $SD = 1.0$), who performed well on the task. In the Hebrew groups, a significant difference was found between the TD-HE group

and the ASD-HE group ($U(30) = 56.5, p < .0001$). In the three French groups, the ASD-FR did not differ from the DLD-FR group ($U(40) = 240.5, p = 0.28$), but performed significantly below the TD-FR group ($U(32) = 209.5, p < .001$).

Unexpected responses

To explore the underlying deficit that was causing low rates of perfect *wh*-questions in each of the three clinical groups (ASD-HE, ASD-FR, and DLD-FR), a comparative analysis of unexpected responses was conducted. As explained in the previous section (and detailed in Table 3), there were two general types of unexpected responses: 1) Incorrect *wh*-questions (morphosyntactically or lexically deviant *wh*-questions) and 2) Pragmatically inappropriate responses (productions resulting from a lack of contextual understanding, which rendered syntactic analysis impossible). We look at these two types of unexpected responses in turn.

Types and rates of Incorrect wh-questions

In the French data, if we only look at the total rate of Incorrect *wh*-questions, the two clinical groups seem, *prima facie*, similar. The ASD and the DLD groups did not differ in the rates of Incorrect *wh*-questions they produced (ASD-FR/DLD-FR: $U(40) = 241, p = .71$), and each of these groups produced more Incorrect *wh*-questions than the TD-FR group (ASD-FR/TD-FR: $U(32) = 63.5, p = .028$; DLD-FR/TD-FR: $U(32) = 49.0, p = .006$, which remains significant under FDR correction). However, when we zoomed in on error types, interesting differences were found between the three French groups. In the ASD-FR group, Other *wh*-questions and Animacy errors were the two dominant error types. The Animacy error rate (16%) did not differ from that of the TD-FR (14.1%, $U(32) = 118.5, p = .95$). The ASD-FR group produced more Other *wh*-questions (17.4% of their responses) than two other groups, almost significantly so in comparison with the TD-FR ($U(32) = 76.5, p = .058$, which ceases to be significant once FDR correction is applied), and not reaching significance in comparison with the DLD-FR group ($U(40) = 138, p = .096$).

In the DLD-FR group, in contrast, the dominant error type was Ungrammatical *wh*-questions (25.3% of their responses). A few Animacy errors and Other *wh*-questions were also found in this group. Crucially, the incidence of Ungrammatical *wh*-questions was relatively low in the ASD-FR group ($M = 3.1%$), and significantly lower than this rate in the DLD-FR group ($U(40) = 320.5, p = .001$). It was, however, marginally higher than in the TD-FR group ($M = 0.7%, SD = 1.3; U(32) = 76.5, p = .058$). Finally, lexical verb substitutions were rare overall in the French groups (fewer than 1% errors in all groups, see Table 4).

Table 4. Types and rates of incorrect *wh*-questions in the ASD, DLD, and TD French-speaking and Hebrew-speaking groups.

		Other <i>wh</i> -questions (%)	Ungrammatical <i>wh</i> -questions (%)	Lexical verb substitutions (%)	Animacy errors (%)	Total Incorrect <i>wh</i>-questions (%)
ASD-FR	<i>M (SD)</i>	17.4 (20.5)	3.1 (3.8)	0.4 (1.0)	16.0 (14.6)	36.8 (21.3)
DLD-FR	<i>M (SD)</i>	6.8 (6.5)	25.3 (23.7)	0.8 (1.6)	8.3 (5.8)	41.3 (19.9)
TD-FR	<i>M (SD)</i>	1.6 (2.2)	0.7 (1.3)	0.7 (1.7)	14.1 (9.7)	17.1 (10.8)
ASD-HE	<i>M (SD)</i>	9.4 (8.7)	2.9 (2.7)	4.3 (3.7)	5.6 (6.4)	22.2 (16.4)
TD-HE	<i>M (SD)</i>	0.4 (1.0)	0 (0)	0 (0)	0 (0)	0.4 (1.0)

In the Hebrew data, the ASD-HE group produced a total of 22.2% Incorrect *Wh* questions, a rate that was significantly higher than in the TD-HE group (0.4%, $U(30) = 188$, $p = .0004$). The types of Incorrect *wh*-questions that the ASD-HE produced were: Other *wh*-questions 9.4%; Animacy errors 5.6%; Lexical verb substitutions 4.3%; and Ungrammatical *wh*-questions 2.9%.

The only type of Incorrect *wh*-questions produced by the children in the TD-HE group (which occurred only twice, once for two different children) was Other *wh*-questions, where a subject question was produced instead of an object question or an object question instead of a subject question. Crucially, the two ASD groups, the FR-ASD and HE-ASD produced few ungrammatical questions (ASD-FR 3.1%; ASD-HE 2.9%), and these rates were much lower than in the DLD-FR group (which was 25.3%).

The two minor differences found between the two ASD groups probably resulted from the difference in elicitation methodology in the two languages. The children in the ASD-FR group produced more Animacy errors than the children in the ASD-HE group (ASD-FR 16%; ASD-HE 5.6%). This is probably related to a difference in methodology: the French task had both animate and inanimate object question items, whereas the Hebrew task had only animate items (As shown in Table 2, even the TD-FR made many Animacy errors, 14.1%, whereas the TD-HE made none). Additionally, lexical verb substitutions were rare in both languages, but were more common in the ASD-HE group (4.3%) than in the ASD-FR group (0.4%). This may be related to the fact that the verb was supplied in the lead-in sentence in the French version of the test, whereas the Hebrew version had no lead-in; however, we note that the TD-HE group made no such errors.

Thus, the two main characteristics of the ASD groups, and the main ways in which they differed from the DLD group were the elevated rate of Other *wh*-questions, and the low rates of Ungrammatical *wh*-questions, which were lower than that seen in the DLD group.

Types and rates of pragmatically inappropriate responses

The second type of errors that were found in both the ASD and DLD groups were unexpected productions that were not appropriate to the context. We chose to analyze these Pragmatically inappropriate responses in detail as these errors could be related to the pragmatic deficit in ASD.

General rate of Pragmatically inappropriate responses.

Figure 5 shows the prevalence of Pragmatically inappropriate responses over all responses, in the three clinical groups. The children in both TD groups (TD-FR and TD-HE) made no such errors and so we do not report on their performance in this section.

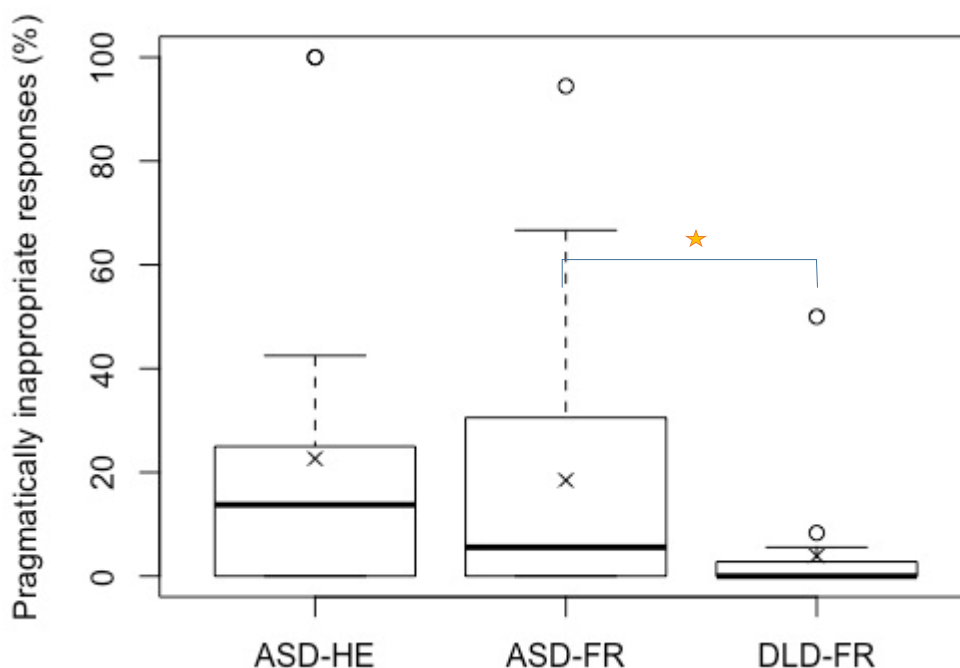


Figure 5: Production rates (%) for Pragmatically inappropriate responses for the three clinical groups, median (bar) and mean (cross). Box and whiskers indicate quartiles based on the mean. Circles indicate outliers. * Indicate a significant difference, $p < .05$

First, it can be noticed that the ASD-FR group, as a group, produced significantly more Pragmatically inappropriate responses than the DLD-FR group (ASD-FR: $M = 18.5\%$, $SD = 26.3$; DLD-FR: $M = 3.9\%$, $SD = 11.1$; $U(40) = 130$, $p = .04$). Interestingly, and despite the methodological differences between the French and the Hebrew tasks, these pragmatically inappropriate responses were found at quite similar rates in the ASD groups in the two languages (ASD-HE: $M = 22.6\%$, $SD = 30.7$; ASD-FR: $M = 18.5\%$, $SD = 26.3$), with high variability in performance in both groups.

The boxplots of Figure 5 show that although the rate of inappropriate responses in both ASD groups was generally very high compared to the DLD-FR group, there were some children who made

no inappropriate responses, while others constantly made them. Participants who produced at least two pragmatically inappropriate responses were counted and supported the results showing that the proportion of inappropriate responses was higher in the children with ASD compared to the children with DLD (ASD-HE: $N = 11/18$, 61%; ASD-FR: $N = 8/20$, 45%; DLD-FR: $N = 2/20$, 10%).

Types of Pragmatically inappropriate responses.

Pragmatically inappropriate responses covered different types of answers that we classified into four categories: Hidden object answers, Yes/No questions, responses Ignoring the event or participant and Other responses (see below for examples of each error type). The distribution of these different types of responses in each ASD group is summed up in Table 5. Percentages were calculated by considering only children that produced at least two Pragmatically inappropriate responses, eight children for French and eleven for Hebrew.

Table 5: Distribution of pragmatically inappropriate answers in ASD-HE and ASD-FR participants producing at least two Pragmatically inappropriate responses

		<i>% of Hidden Object Answers</i>	<i>% of Yes/No questions</i>	<i>% of Ignoring event or participant</i>	<i>% of Other responses</i>
ASD-FR ($n = 8$)	M (SD)	27.00 (27.28)	22.42 (31.86)	45.67 (37.98)	4.45 (6.55)
ASD-HE ($n = 11$)	M (SD)	32.98 (28.98)	1.67 (2.43)	53.89 (29.22)	11.45 (9.28)

Types of inappropriate responses in the ASD-FR group.

The most common error found among Inappropriate responses was 'Ignoring event or participant'. These were responses in which the participant produced a question, but failed to take into account the action presented in the stimulus (and used a verb such as 'do' or 'be' instead) or one (or both) of the participants, as in (5).

(5) Ignoring event or participant - Subject question

Qu'est-ce que tu fais, éléphant? (Child BRL, ASD-FR, 6;10)

'What are you doing, elephant?'

Target: Who is getting you wet?

These particular responses represented about half of the responses of the French-speaking children

with ASD who produced at least two pragmatically inappropriate responses.⁴

Children with ASD also produced a large number of 'Hidden object responses', the second most common type of Inappropriate responses. Hidden object responses were responses in which, instead of producing a target *wh*-question, the participant guessed (either by question or by affirmation) what was behind the white shape (6a); or the participant expressed knowledge about the hidden object or character (6b).

(6) a. Guessing hidden object – Object *what* question

Chien, tu mords l'os? (Child ARF, ASD-FR, 12;9)

'Dog, are you biting a bone?'

Target: What are you biting?

b. Knowledge of hidden object - Object *who* question

C'est l'éléphant qui mouille le lion. (Child FRA, ASD-FR, 6;2)

'It's the elephant that is getting the lion wet.'

Target: Who are you getting wet?

Yes/No questions, as in (7), were also found, at a moderate rate. These were cases where the participant produced a yes/no question instead of the target subject/object *wh*-question.

(7) Yes/No question - Subject question

Est-ce que quelqu'un te lave? (Child MAD, ASD-FR, 9;2)

'Is someone washing you?'

Target: Who is washing you?

Finally, other types of responses, less common, were gathered into an 'Other responses' category. All of these responses involved some kind of repetition: repetition of part of the stimulus or verbatim repetition of a previous response (8).

(8) Repetition of the stimulus - Subject question

Demande au chien qui le pousse. (Child JOC, ASD-FR, 6;4)

'Ask the dog who is pushing him.'

Target: Who is pushing you?

⁴ These productions were also the only ones found in the two DLD-FR children producing at least two inappropriate responses

Table 6 presents the categories of Inappropriate responses produced by each of the eight ASD-FR children having produced two or more such responses. It is striking that the dominant inappropriate response varied from child to child and that for most children the dominant response was (almost) the sole type of inappropriate response produced (see, for example, BRL, WAE, JUF, and ARF).

Table 6. Individual ASD-FR results for types of inappropriate responses (dominant responses in boldface).

<i>Child</i>	<i>Age (months)</i>	<i>Number of inappropriate responses (/36)</i>	<i>Ignoring event or participant (%)</i>	<i>Hidden object responses (%)</i>	<i>Yes/No questions (%)</i>	<i>Other Responses (%)</i>
JOC	76	34	0	73	6	20
BRL	83	13	92	0	8	0
WAE	83	24	0	0	100	0
BRR	86	9	78	11	0	11
JUF	86	6	100	0	0	0
MAD	110	17	12	0	88	0
SEG	112	3	67	33	0	0
ROD	122	8	62	25	0	12
ARF	153	14	0	100	0	0

Types of inappropriate responses in the ASD-HE group.

Table 7 shows the distribution of the error types found in the 11 ASD-HE participants who gave at least two inappropriate responses. As was seen in the ASD-FR group, some participants tended to use one specific type of inappropriate response.

Table 7. Individual ASD- HE results for types of inappropriate responses (predominant responses in boldface).

<i>Child</i>	<i>Age (months)</i>	<i>N Inappropriate responses (/40)</i>	<i>Ignoring event or participant (%)</i>	<i>Hidden Object (%)</i>	<i>Yes/ No questions (%)</i>	<i>Other (%)</i>
OG	115	17	35	29	6	18
EL	121	40	35	48	3	0
OW	123	13	38	54	0	0
TZ	127	8	0	100	0	0
TR	136	40	0	90	10	0
LO	140	7	86	14	0	0
AC	142	5	100	0	0	0
TW	147	4	75	0	0	25
MS	194	6	67	17	0	17
NZ	219	10	90	0	0	10
AG	222	9	67	11	0	0

The two most dominant inappropriate responses in the ASD-HE group were Ignoring event or participant, as in (9), and Hidden Object responses, as in (10). Several participants also produced yes/no questions instead of the target subject/object question as in (11). Other responses, as in (12), included some kind of repetition, repetition of part of the stimuli or verbatim repetition of previous responses.

- (9) Ignoring event or participant - Subject question

Mi martiv? (Child MS, ASD-HE, 16;1)

Who wet?

'Who is wetting?'

Target: Who is spraying the man with water?

- (10) Hidden Object response - Object question

Ha'im ha-xatul melakek et ha-kelev? (Child TR, ASD-HE, 11;3)

whether the-cat licking ACC the-dog?

'Is the cat licking the dog?'

Target: Who is the cat licking?

- (11) Yes/ No response - Object question

Ha'im ha-arie memale et ha-ambatia? (Child TR, ASD-HE, 11;3)

whether the-lion filling ACC the-bath

'Is the lion filling the bath?'

Target: Who is the lion bathing?

- (12) Other responses- Object question

Mi menagev? (Child MS, ASD-HE, 16;16)

who drying?

'Who is drying?'

Target: Who is the prince tucking in?

(this response is perseverative: For three consecutive stimuli, the participant's productions included the verb *menagev* 'drying')

Once again, some similarities were found between the two ASD groups when zooming in on pragmatically inappropriate responses. In both groups, the most predominant error among inappropriate responses was Ignoring event or participant. These particular responses represented

about half responses of both the Hebrew-speaking and the French-speaking children with ASD who produced at least two pragmatically inappropriate responses. The two ASD groups also resembled each other by the large number of Hidden object answers, the second most common type of inappropriate responses. The French-speaking children produced Yes/No questions more often than did the Hebrew-speaking children, who in turn produced Other responses more frequently.

Correlation analyses

The results in the preceding section showed that performance of children with ASD on production of *wh*-questions is characterized, in both languages, by high rates of unexpected responses (either pragmatically inappropriate or linguistically incorrect) and considerable variability. In order to better understand these results, a series of correlation analyses were run on each of the two ASD groups, involving age, NVIQ, and independent morphosyntactic measures (a standardized sentence completion task for French and a sentence repetition task for Hebrew).

For French, no significant correlations were found between age and any of the production rates of our three main responses: Perfect *wh*-questions ($r_s = .28, p = .24$), Pragmatically inappropriate responses ($r_s = -.31, p = .19$), and Incorrect *wh*-questions ($r_s = .24, p = .30$). The same held true for NVIQ, measured by RPM scores (Perfect *wh*-questions: $r_s = .29, p = .22$; Pragmatically inappropriate responses: $r_s = -.28, p = .23$; Incorrect *wh*-questions: $r_s = .09, p = .71$). In contrast, not surprisingly, expressive morphosyntax scores were significantly correlated with responses hinging on morphosyntax: the rates of Perfect *wh*-questions ($r_s = .52, p = .02$) and Incorrect *wh*-questions ($r_s = -.62, p = .004$), but not with the rate of Pragmatically inappropriate responses ($r_s = -.22, p = .35$).

For Hebrew, age was significantly correlated with the production rates of Perfect *wh*-questions ($r_s = .53, p = .03$) and Pragmatically inappropriate responses ($r_s = -.55, p = .02$), but not with the rate of Incorrect *wh*-questions ($r_s = -.049, p = .85$). Concerning nonverbal reasoning, we found that scores for the Nonverbal Picture Association task were not correlated with any response type (Perfect *wh*-questions: $r_s = 0.14, p = .56$; Pragmatically inappropriate responses: $r_s = 0.05, p = .86$; Incorrect *wh*-questions: $r_s = -0.22, p = .36$). The morphosyntactic measure, Sentence Repetition, was significantly correlated with the rate of Incorrect *wh*-questions ($r_s = -0.59, p = .01$), but not with the rates of Perfect *wh*-questions ($r_s = 0.21, p = .38$) or Pragmatically inappropriate responses ($r_s = 0.12, p = .61$).

Summary

The results for French showed that a group of 20 French-speaking children with ASD, aged 6 to 12 and of widely varying nonverbal ability levels, resembled age-matched children with DLD in their rates of production of expected *wh*-questions and in the fact that they performed significantly below TD children aged 6. ASD/DLD similarities, but also differences, emerged when unexpected responses were analyzed and categorized into those which were inappropriate to the item prompts and pictures, and those which were not. The ASD group, once again, resembled the DLD group for rates of incorrect *wh*-questions (and both groups performed significantly below the TD 6-year-olds). The ASD group, however, produced strikingly and significantly higher rates of inappropriate responses compared to the DLD group, where such responses were made by only two children, and to the TD children, who did not produce any responses of this type.

The results for Hebrew showed that a group of 18 Hebrew-speaking children with ASD, aged 9 to 18, with no known intellectual disability, performed much below younger TD children (aged 7-9 years) in their rates of production of expected *wh*-questions. Although the TD children were significantly younger than most of the ASD children, they completed the task with ceiling scores. Particularly striking among the unexpected *wh*-questions that were found were inappropriate responses, which were produced by more than half of the participants with ASD, but not at all in the TD group.

Despite the differences between the properties of French and Hebrew, the differences in methodology, and the differences in participant characteristics in the two studies, our comparative results showed striking resemblance. In both languages, the children with ASD produced fewer correct *wh*-questions compared to TD children, and a fine-grained analysis of error types revealed a substantial number of pragmatic errors used as responses that were inappropriate for specific test items or for the general testing context. Such pragmatic errors, found in the two ASD groups, were not seen in either the TD or the DLD groups.

General Discussion

This study investigated *wh*-question production in children with ASD across two different languages in order to explore the nature of observed difficulties in *wh*-question production in this population. Two studies examined, respectively, French-speaking and Hebrew-speaking children with ASD in order to determine how frequently these children produced expected *wh*-questions, ones that were context-appropriate and constituted syntactically well-formed *wh*-questions. Unexpected

productions were subjected to further analysis in order to determine the nature of the errors—were these context-appropriate, and if so, did they contain morphosyntactic or lexical errors?

Regarding the first question, how frequently children with ASD produce context-appropriate and syntactically well-formed *wh*-questions, our results indicate that both the French-speaking and the Hebrew-speaking children with ASD produced (similarly) low rates of Perfect *wh*-questions (ASD-FR – 41%; ASD-HE – 55%), which were context-appropriate and syntactically well-formed, significantly lower than those found in the younger TD children groups. These similarly low rates were found despite the fact that the syntax of *wh*-questions differs in the two languages tested. Furthermore, the two French-speaking clinical groups, ASD-FR and DLD-FR, showed similar rates of perfect *wh*-questions, in line with previous studies (Demouy et al., 2011; Sukenik & Friedmann, 2018). Response variability in the clinical groups was quite high and could be attributed to well-known variation in severity of language impairment. Had we concluded our study at this point, looking only at rates for expected and syntactically well-formed *wh*-questions, the conclusion would have been that children with ASD resemble children with DLD in their ability to produce *wh*-questions, a conclusion that meshes with previous studies finding similarities in syntactic (dis)abilities between these two populations.

However, a crucial conclusion emerges from the second research question, exploring the nature of unexpected *wh*-questions via error patterns. We asked whether children’s difficulty in producing *wh*-question was due to syntactic/lexical or pragmatic shortcomings. Syntactic/lexical errors in the two ASD groups were not negligible (22% of all responses in the ASD-HE group and 37% in the ASD-FR group, which was not significantly different from the 41% found in the DLD-FR group). This result would appear to be in harmony with the conclusion reached above regarding over-all *wh*-question abilities: the participants with ASD appeared to resemble the children with DLD in the frequency of erroneous *wh*-questions, at least at group level.

However, ASD/DLD differences were evident when error typology was considered. The most common error, by far, in the DLD group was Ungrammatical *wh*-questions (25.3%), an error that was much less frequent in the two ASD groups (ASD-FR 3.1%; ASD-HE 2.9%). In the ASD groups, on the other hand, it was Other *wh*-questions (*wh*-questions of a different type—subject instead of object question, *where*-question instead of participant-question, etc.) that was the most frequent error type (ASD-FR 17.4%; ASD-HE 9.4%). These errors were also found in the DLD group, but much less frequently.

The difference in the relative frequency of different error types between the ASD and DLD groups suggests that many errors observed in the ASD group are not really syntactic/lexical in nature. Regarding the questions that were coded as “Other *wh*-questions”, there may be more than one underlying cause for these responses. On the one hand, such responses could reflect difficulties in switching (an ASD trait related to social and executive functioning, Akbar et al., 2013; DiCriscio et al., 2019; Reed & McCarthy, 2012) from subject to object questions in the task. Such difficulty in switching between question structures could create perseveration of a particular question form (use of the same *wh*-word/phrase/structure regardless of the stimuli presented). For example, in the ASD-HE group, several participants produced subject questions throughout the task. These responses could have resulted from a syntactic difficulty, as subject questions in Hebrew are less complex than object questions (the *wh*-movement that derives them does not cross any other argument). However, at least in some cases there are indications that substitution actually resulted from perseveration. One type of perseveration was of the *wh* element itself: when the target question was an object question, some participants produced the *wh*-word *mi* ‘who’, which is used in subject questions, instead of *et mi* (accusative marker + *wh* element), but kept the rest of the structure as an object question structure (*mi ha-kelev noshex*). Here, the syntactic structure was appropriate for the target question, but the perseveration of the *wh* element created an ungrammatical question. Another type of perseveration was the production of a consecutive sequence of object questions, the more complex structure, regardless of the target question types. In the ASD-FR group several participants used the phrase *qu’est-ce-que* ‘what + question morpheme *est-ce que*’, a fixed expression only compatible with an object *wh*-question, repeatedly, and therefore produced non-target ‘what’ questions in place of ‘who’ questions, non-target object questions in place of subject questions, as well as target ‘what’ object questions. Such production of sequences of object questions, creating non-target object questions produced instead of subject questions, points to a perseveration source for the production of incorrection question type, rather than a syntactic complexity source.

The other type of unexpected responses were Inappropriate responses, responses that were pragmatically not appropriate to the stimulus prompt and picture. These types of responses were quite frequent in both the French-speaking and the Hebrew-speaking ASD groups. In contrast, they were nonexistent in the corresponding TD groups of younger children, and marginal in the DLD-FR group. Nave, Szterman, and Friedmann (2009) and Friedmann and Szterman (2011) used this same task to elicit *wh*-questions in Hebrew-speaking children with hearing impairment (HI). The HI

children produced many syntactic errors, especially in target object questions, but no inappropriate responses of any kind were observed. In other words, despite the fact that the structure of *wh*-questions is different in French and Hebrew, that the task used in the two languages was structured somewhat differently, and that the children with ASD in the French study were younger than those in the Hebrew study, the two ASD groups had similarly high rates of pragmatic errors. Moreover, the types of pragmatic errors found in each of the ASD groups were the same. Previous studies found children with ASD to have trouble understanding the social-communicative purposes of questions, i.e., that questions are used to elicit information from an interlocutor (Koegel et al., 2014; Matson et al., 1996; Weiss & Harris, 2001). The task used in our study required children to understand that they needed to ask a question in order to find out who/what the hidden entity was. Many children with ASD had trouble mapping the event and the character(s) depicted in the picture onto a *wh*-question. Although the French task included a lead-in sentence that named the event and the participant in the picture, ignoring event and/or participant was the dominant pragmatic error in the ASD-FR group, as well as in the ASD-HE group, where the task did not include a lead-in. Two expressions of this basic difficulty were yes-no questions/statements about the identity of one of the participants or about the shape hiding the second participant (Yes-no question errors and Hidden object errors). Errors in the remaining pragmatic error category, which grouped non-responses, echolalic responses, multiple *wh*-question responses and perseverations from previous stimuli also seemed to show that the participants had trouble adapting to the task context.

In order to try to explain the errors found in the ASD groups, correlation analyses were conducted between the different response types and the background measures of each of the language groups. In both ASD groups the non-verbal measures (RPM for French and picture association for Hebrew) were not found to be correlated to any of the response types. This result highlights the importance of testing more than one cognitive ability (i.e. nonverbal abilities as well as language abilities) as these may be functioning at different levels. In both ASD groups, an independent morphosyntactic measure was correlated to the rate of Incorrect *wh*-questions (in the French ASD group also to the rate of perfect *wh*-questions), but not to the rate of Pragmatically inappropriate responses. These findings support the hypothesis that at least some of the difficulties that children with ASD demonstrate while producing *wh*-questions may stem from a syntactic deficit.

A negative correlation between Pragmatically deviant responses and age was found in the ASD-HE group, but not in the ASD-FR group. Not only was the age range wider in the ASD-HE group

compared to the ASD-FR group (9-18 versus 6-12), many of Hebrew-speaking children were simply older. Most of the children with ASD in both studies had benefited from speech-language therapy. In therapy, children with ASD are explicitly coached on pragmatic skills. Since such explicit learning is a long process, it is not surprising that older children would be more likely to have learned how to adapt to a testing situation such as that inherent in the *wh*-question production task reported on here, and thus should have been more able to produce pragmatically appropriate responses. The fact that a (negative) correlation was not found between age and Incorrect *wh*-questions would seem to indicate that structural language abilities have not improved with age, perhaps because therapy has mainly targeted pragmatics, rather than, for example, syntax.

This study sought to explore and understand the underlying mechanisms that were causing low performance on *wh*-production in children with ASD. We used a structured elicited production task with the aim of eliciting many more questions than usually seen in the free speech of children with ASD. Applying a coding scheme based on whether or not the child's production consisted of a *wh*-question which was relevant to the context, and, if so, whether or not it was a grammatical question of the expected type, we found that both syntactic and pragmatic deficits seemed to contribute to low performance on *wh*-question production. Finally, although every effort was made to provide a motivated classification of errors as either fundamentally pragmatic or syntactic/lexical in nature, some errors could have both of these sources. Nevertheless, the complete analysis of response types indicates that the difficulty in *wh*-questions has to include a pragmatic component and cannot be explained entirely as a syntactic deficit. This finding has broad implications for future research but also for therapy and educational settings.

In order for children with ASD to receive appropriately targeted therapy it would seem to be fundamental that they be specifically evaluated in different language domains, including pragmatics and syntax. Specific evaluation of the latter entails filtering out potential pragmatic effects as much as possible. We believe that there would be much to be gained in future studies exploring syntactic abilities of children with ASD by 'minimalizing' tasks in this way. Such methodological precautions would be particularly useful for comparative studies seeking to test the similarities and differences between ASD and DLD and look more closely at the underlying linguistic mechanisms that are affecting linguistic comprehension and production in these populations. How could the morphosyntax of question production be assessed in children with ASD independently of pragmatics? Several recent studies tested individuals with ASD (Eilon, 2013; Eilon et al., 2014; Silleresi, 2018, 2020; Sukenik, 2017; Yarkoni-Kaplan, 2014; Zaidenberg, 2015; Zaidenberg et al., 2019) using sentence

repetition tasks (wherein the participant heard a sentence and was asked to repeat it exactly as (s)he heard it) that included *wh*-questions, thus testing production in the absence of a communicative context. The production of *wh*-questions elicited in this manner did not give rise to the number of pragmatically deviant responses we have reported on here, an encouraging result, which points to a possible way to examine the syntactic aspects of *wh*-questions, as well as other types of complex sentences, with far fewer pragmatic requirements. When considering therapeutic settings, the results of this study highlight the need to teach children with ASD both the syntactic aspects of questions, and the pragmatic aspects and the social usage of questions.

We fully acknowledge the limitations of this study, which stemmed from the decision to pool and subject to further analysis data from two separate studies conducted in different countries. This meant that the available data had already been collected and further data collection was not possible and thus background measures were not identical, and the elicitation method differed, making statistical comparison of the two groups impossible. We also emphasize that children were not recruited in the same types of structures in the two countries, resulting in two very different convenience samples, one from a clinical setting and one from an educational setting. We believe that the similarity in results is all the more remarkable, though controlling these elements would clearly be preferable in future cross-linguistic/cultural studies of language in ASD.

In summary, the comparison of *wh*-question production in two different languages, which utilize different syntactic structures, found similar syntactic and pragmatic errors in both ASD groups. Some error types were argued to most likely be both pragmatic and syntactic in nature, making it extremely difficult to identify the nature of production difficulties through tasks of this type in children with ASD. Our results highlight the importance of creating research tasks that test different linguistic functions independently and strengthen the need for conducting fine grained error analysis as previously suggested (Sukenik & Friedmann, 2018). If only correct *wh*-question rates had been compared, children with ASD and DLD would have been found to have similar performance. Finally, this study has provided a glimpse into the shared autism traits which seem to be indifferent to the specific culture of the child and seem to affect all children with this disorder, and which make evaluation of language abilities of these children particularly challenging.

References

- Adams, C. (1990). Syntactic comprehension in children with expressive language impairment. *International Journal of Language & Communication Disorders, 25*(2), 149-171.
- Akbar, M., Loomis, R., & Paul, R. (2013). The interplay of language on executive functions in children with ASD. *Research in Autism Spectrum Disorders, 7*(3), 494-501.
- Armon-Lotem, S. (2008). The interaction between question formation and verbal morphology in the acquisition of Hebrew – a Minimalist perspective. In S. Armon-Lotem, G. Danon & S. Rothstein (Eds.), *Current Issues in Generative Hebrew Linguistics* (pp. 223-243). Benjamins.
- Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a “Theory of mind”? *Cognition, 21*(1), 37-46.
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological), 57*(1), 289-300.
- Biran, M. & Ruigendijk, E. (2015). Do case and gender information assist sentence comprehension and repetition for German and Hebrew-speaking children? *Lingua, 164*, 215-238. doi: <http://dx.doi.org/10.1016/j.lingua.2015.06.012>
- Brock, J., Sukenik, N., & Friedmann, N. (2017). Individual differences in autistic children's homograph reading: Evidence from Hebrew. *Autism & Developmental Language Impairments, 2*, 1-10. doi: 10.1177/2396941517714945
- Condouris, K., Meyer, E., & Tager-Flusberg, H. (2003). The relationship between standardized measures of language and measures of spontaneous speech in children with autism. *American Journal of Speech-Language Pathology, 12*(3), 349-358.
- Demouy, J., Plaza, M., Xavier, J., Ringeval, F., Chetouani, M., Perisse, D., & Robel, L. (2011). Differential language markers of pathology in autism, pervasive developmental disorder not otherwise specified and specific language impairment. *Research in Autism Spectrum Disorders, 5*(4), 1402-1412.
- de Villiers, P. (2004). Assessing pragmatic skills in elicited production. *Seminars in Speech and Language, 25* (1), 57-71.
- DiCriscio, A. S., Hu, Y., & Troiani, V. (2019). Brief Report: Pupillometry, Visual Perception, and ASD Features in a Task-Switching Paradigm. *Journal of autism and developmental disorders, 49*(12), 5086-5099.
- Durrleman, S., & Delage, H. (2016). Autism spectrum disorder and specific language impairment: Overlaps in syntactic profiles. *Language Acquisition, 23*(4), 361-386.
- Durrleman, S., Marinis, T., & Franck, J. (2016). Syntactic complexity in the comprehension of wh-questions and relative clauses in typical language development and autism. *Applied Psycholinguistics, 37*(6), 1501-1527.
- Durrleman, S., Delage, H., Prévost, P., & Tuller, L. (2017). The comprehension of passives in autism spectrum disorder. *Glossa, 2*, 1-30.
- Eigsti, I. M., Bennetto, L., & Dadlani, M. B. (2007). Beyond pragmatics: Morphosyntactic development in autism. *Journal of Autism and Developmental Disorders, 37*(6), 1007-1023.
- Eilon, I. (2013). *The Influence of the Impairment in Theory of Mind on the Linguistic Abilities of Children with ASD* (Unpublished Master's thesis). Tel Aviv University, Tel Aviv.
- Eilon, I., Balaban, N., & Friedmann, N. (2014). *Theory of Mind impairment and linguistic abilities in autism*. Presented at the International Neuropsychological Society Meeting, Jerusalem.
- Eigsti, I.-M., & Schuh, J. M. (2017). Language acquisition in ASD: Beyond standardized language measures. In L. R. Naigles (Ed.), *Language and the human lifespan series. Innovative investigations of language in autism spectrum disorder* (p. 183-200). American Psychological Association; Walter de Gruyter. <https://doi.org/10.1037/15964-010>
- Friedmann, N. (1998). *BAFLA: Friedmann Battery for Agrammatism*. Tel Aviv: Tel Aviv University.
- Friedmann, N. (2002). Question production in agrammatism: The tree pruning hypothesis. *Brain and Language, 80*(2), 160-187.
- Friedmann, N., Belletti, A., & Rizzi, L. (2009). Relativized relatives: Types of intervention in the acquisition of A-bar dependencies. *Lingua, 119*(1), 67-88.

- 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.
- Friedmann, N., & Novogrodsky, R. (2008). Subtypes of SLI: SySLI, PhoSLI, LeSLI, and PraSLI. In A. Gavarró & M. J. Freitas (Eds.), *Language acquisition and development* (pp. 205-217). Cambridge Scholars Publishing.
- Friedmann, N., & Novogrodsky, R. (2011). Which questions are most difficult to understand? The comprehension of Wh questions in three subtypes of SLI. *Lingua*, 121(3), 367-382.
- Friedmann, N., & Szterman, R. (2011). The comprehension and production of Wh-questions in deaf and hard-of-hearing children. *Journal of Deaf Studies and Deaf Education*, 16(2), 212-235.
- Friedmann, N., Yachini, M., & Szterman, R. (2015). Relatively easy relatives: Children with syntactic SLI avoid intervention. In E. Di Domenico, C. Hamann, and S. Matteini (Eds.), *Structures, strategies and beyond. Studies in honour of Adriana Belletti* (pp. 303-320). Amsterdam: John Benjamins.
- Gavarró, A., & Heshmati, Y. (2014). An investigation on the comprehension of Persian passives in typical development and autism. *Catalan Journal of Linguistics*, 13, 79-98.
- 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(2), 109-123.
- Håkansson, G., & Hansson, K. (2000). Comprehension and production of relative clauses: A comparison between Swedish impaired and unimpaired children. *Journal of Child Language*, 27(2), 313-333.
- Hamann, C. (2006). Speculations about early syntax: The production of wh-questions by normally developing French children and French children with SLI. *Catalan Journal of Linguistics*, 5(1), 143-189.
- Hamann, C., & Tuller, L. (2015). Intervention effects in the spontaneous production of relative clauses in (a) typical language development of French children and adolescents. In E. Di Domenico, C. Hamann, and S. Matteini (Eds.), *Structures, strategies and beyond. Studies in honour of Adriana Belletti* (pp. 321-342). Benjamins.
- Hoff, E. (2013). *Language development*. Cengage Learning.
- Jakubowicz, C. (2011). Measuring derivational complexity: New evidence from typically developing and SLI learners of L1 French. *Lingua*, 121(3), 339-351.
- Jyotishi, M., Fein, D., & Naigles, L. (2017). Investigating the grammatical and pragmatic origins of wh-questions in children with autism spectrum disorders. *Frontiers in Psychology*, 8, 319.
- Khomsí, A., Khomsí, J., Pasquet, F., & Parbeau-Guéno, A. (2007). Bilan Informatisé de Langage Oral au cycle 3 et au Collège (BILO-3C). Paris, France: Éditions du Centre de Psychologie Appliquée.
- Kjelgaard, M. M., & Tager-Flusberg, H. (2001). An investigation of language impairment in autism: Implications for genetic subgroups. *Language and Cognitive Processes*, 16(2-3), 287-308.
- Koegel, R. L., Bradshaw, J. L., Ashbaugh, K., & Koegel, L. K. (2014). Improving question-asking initiations in young children with autism using pivotal response treatment. *Journal of Autism and Developmental Disorders*, 44(4), 816-827.
- Levy, H., & Friedmann, N. (2009). Treatment of syntactic movement in syntactic SLI: A case study. *First Language*, 29(1), 15-49.
- Lombardo, M. V., Lai, M. C., & Baron-Cohen, S. (2019). Big data approaches to decomposing heterogeneity across the autism spectrum. *Molecular Psychiatry*, 24(10), 1435-1450.
- Lord, C., Rutter, M., Goode, S., Heemsbergen, J., Jordan, H., Mawhood, L., & Schopler, E. (1989). Autism diagnostic observation schedule: A standardized observation of communicative and social behavior. *Journal of Autism and Developmental Disorders*, 19(2), 185-212.
- Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism diagnostic interview-revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, 24, 659-685.
- Masi, A., DeMayo, M. M., Glozier, N., & Guastella, A. J. (2017). An overview of autism spectrum disorder, heterogeneity and treatment options. *Neuroscience Bulletin*, 33(2), 183-193.
- Matson, J., Benavidez, D., Stabinsky Compton, L., Paclawsky, T., & Baglio, C. (1996). Behavioral treatment of autistic persons: A review of research from 1980 to the present. *Research in Developmental Disabilities*, 17, 443-465.

- McGregor, K. K., Berns, A. J., Owen, A. J., Michels, S. A., Duff, D., Bahnsen, A. J., & Lloyd, M. (2012). Associations between syntax and the lexicon among children with or without ASD and language impairment. *Journal of Autism and Developmental Disorders*, 42(1), 35–47.
- Modyanova, N., Perovic, A., & Wexler, K. (2017). Grammar is differentially impaired in subgroups of autism spectrum disorders: evidence from an investigation of tense marking and morphosyntax. *Frontiers in Psychology*, 8, 320.
- Naigles, L., Cheng, M., Rattanasone, N. X., Tek, S., Khetrapal, N., Fein, D., & Demuth, K. (2016). “You’re telling me!” The prevalence and predictors of pronoun reversals in children with autism spectrum disorders and typical development. *Research in Autism Spectrum Disorders*, 27, 11-20.
- Novogrodsky, R., & Friedmann, N. (2006). The production of relative clauses in syntactic SLI: A window to the nature of the impairment. *Advances in Speech Language Pathology*, 8(4), 364-375.
- Park, J. (2016). *Wh-question acquisition in Korean: Typically developing children and children with autism spectrum disorder* (Doctoral Dissertation). University of Connecticut. <https://opencommons.uconn.edu/dissertations/1309/>
- Prévost, P., Tuller, L., Scheidnes, M., Ferré, S., & Haiden, M. (2010). Computational Complexity effects in the acquisition of wh-questions in L2 child French. In L. Dominguez & P. Guijarres-Fuentes (Eds.), *New directions in language acquisition: Romance languages in the generative perspective* (pp. 415-443). Cambridge Scholars Publishing.
- Prévost, P., Tuller, L., Zebib, R., Barthez, M. A., Malvy, J., & Bonnet-Brilhault, F. (2018). Pragmatic versus structural difficulties in the production of pronominal clitics in French-speaking children with autism spectrum disorder. *Autism & Developmental Language Impairments*, 3, 1-17.
- Rapin, I., & Dunn, M. (2003). Update on the language disorders of individuals on the autistic spectrum. *Brain and Development*, 25(3), 166-172.
- Reed, P., & McCarthy, J. (2012). Cross-modal attention-switching is impaired in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 42(6), 947-953.
- Reznick, J., & Friedmann, N. (2017). On the order of acquisition of various types of movement in Hebrew. *Language and Brain*, 12, 109-168. (in Hebrew).
- Riches, N. G., Loucas, T., Baird, G., Charman, T., & Simonoff, E. (2010). Sentence repetition in adolescents with specific language impairments and autism: An investigation of complex syntax. *International Journal of Language & Communication Disorders*, 45(1), 47-60.
- Roberts, J. A., Rice, M. L., & Tager-Flusberg, H. (2004). Tense marking in children with autism. *Applied Psycholinguistics*, 25(3), 429-448.
- Schuele, C. M., & Tolbert, L. (2001). Omissions of obligatory relative markers in children with specific language impairment. *Clinical Linguistics & Phonetics*, 15(4), 257-274.
- Seidl, A., Hollich, G., & Jusczyk, P. W. (2003). Early understanding of subject and object wh-questions. *Infancy*, 4(3), 423-436.
- Seymour, H. N., Roeper, T. W., & De Villiers, J. G. (2003). *DELV: Diagnostic Evaluation of Language Variation: Norm Referenced*. Pearson.
- Silleresi, S. (2018). *Structural language and nonverbal ability profiles in monolingual and bilingual children with ASD*. Doctoral dissertation, University of Tours. <https://www.theses.fr/2018TOUR2027>.
- Silleresi, S., Prévost, P., Zebib, R., Bonnet-Brilhault, F., Conte, D., & Tuller, L. (2020). Identifying language and cognitive profiles in children with ASD via a cluster analysis exploration: Implications for the new ICD-11. *Autism Research*, 17, 1155-1167.
- Stavrakaki, S. (2001). Comprehension of reversible relative clauses in specifically language impaired and normally developing Greek children. *Brain and Language*, 77(3), 419-431.
- Su, Y. E., Jin, Y., Wan, G. B., Zhang, J. S., & Su, L. Y. (2014). Interpretation of wh-words in Mandarin-speaking high-functioning children with autism spectrum disorders. *Research in Autism Spectrum Disorders*, 8(10), 1364-1372.
- Sukenik, N. (2012). *The relationship between comprehension and reading at a word, sentence and paragraph level in children with ASD*. [Unpublished Master's thesis]. Tel Aviv University.

- Sukenik, N. (2017). *Sémantique lexicale et profils langagiers d'enfants avec autisme de langue hébraïque* (Doctoral dissertation). Université François-Rabelais de Tours. Retrieved from <https://www.theses.fr/2017TOUR2024>
- Sukenik, N., & Friedmann, N. (2018). ASD is not DLI: Individuals with autism and individuals with syntactic DLI show similar performance level in syntactic tasks, but different error patterns. *Frontiers in Psychology, 9*, 279.
- Tek, S., Mesite, L., Fein, D., & Naigles, L. (2014). Longitudinal analyses of expressive language development reveal two distinct language profiles among young children with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 44*(1), 75-89.
- Terzi, A., Marinis, T., Kotsopoulou, A., & Francis, K. (2014). Grammatical abilities of Greek-speaking children with autism. *Language Acquisition, 21*(1), 4-44.
- Tuller, L., Ferré, S., Prévost, P., Barthez, M. A., Malvy, J., & Bonnet-Brilhault, F. (2017). The effect of computational complexity on the acquisition of French by children with ASD. In L. Naigles (Ed.), *Innovative Investigations of Language in Autism Spectrum Disorder* (pp. 115-140). de Gruyter.
- van der Lely, H. K., & Harris, M. (1990). Comprehension of reversible sentences in specifically language-impaired children. *Journal of Speech and Hearing Disorders, 55*(1), 101-117.
- Weiss, M. J., & Harris, S. L. (2001). Teaching social skills to people with autism. *Behavior Modification, 25*(5), 785-802.
- Williams, D., Botting, N., & Boucher, J. (2008). Language in autism and specific language impairment: Where are the links?. *Psychological Bulletin, 134*(6), 944-963.
- Yarkoni-Kaplan, T. (2014). *Comprehension and use of pronouns in children with Specific Language Impairment (SLI) and children with autistic spectrum disorders (ASD)*. (Unpublished Master's thesis). Tel Aviv University, Tel Aviv.
- Zaidenberg, H. (2015). *Prosodic deficit in the perception of focus: Evidence from Hebrew-speaking individuals with Asperger Syndrome*. (Unpublished Master's thesis). Tel Aviv University, Tel Aviv.
- Zaidenberg, H., Friedmann, N., & Bat-El, O. (2019). *On the nature of prosody perception deficits in ASD: Evidence from Hebrew speaking high-functioning adults*. Presented at the European conference on language disorders (EUCLDIS) meeting, Tel Aviv.
- Zebib, R., Tuller, L., Prevost, P., & Morin, E. (2013). Formal language impairment in French-speaking children with ASD: A comparative ASD/SLI study. In S. Stavrakaki, M. Lalioti & P. Konstantinopoulou (Eds.), *Advances in language acquisition* (pp. 472-484). Cambridge Scholars Publishing.