Syntactic complexity in the comprehension of wh-questions and relative clauses in typical language development and autism

Article

Accepted Version

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

Syntactic complexity in the comprehension of wh-questions and relative clauses in

typical language development and autism

Stephanie Durrleman[1], Theodoros Marinis[2], Julie Franck[1]

[1]University of Geneva, [2]University of Reading

Running Head: Syntactic complexity in TD and ASD children

Abstract

This study investigates effects of syntactic complexity operationalised in terms of movement, intervention and (NP) feature similarity in the development of A' dependencies in 4-, 6-, and 8-year old typically developing (TD) French children and children with Autism Spectrum Disorders (ASD). Children completed an off-line comprehension task testing eight syntactic structures classified in four levels of complexity: Level 0: No Movement; Level 1: Movement without (configurational) Intervention; Level 2: Movement with Intervention from an element which is maximally different or featurally 'disjoint' (mismatched in both lexical NP restriction and number); Level 3: Movement with Intervention from an element similar in one feature or featurally 'intersecting' (matched in lexical NP restriction, mismatched in number). The results show that syntactic complexity affects TD children across the three age groups, but also indicate developmental differences between these groups. Movement affected all three groups in a similar way, but intervention effects in intersection cases were stronger in younger than older children, with NP feature similarity affecting only 4-year olds. Complexity effects created by the similarity in lexical restriction of an intervener thus appear to be overcome early in development, arguably thanks to other differences of this intervener (which was mismatched in number). Children with ASD performed less well than the TD children although they were matched on non-verbal reasoning. Overall, syntactic complexity affected their performance in a similar way as in their TD controls, but their performance correlated with non-verbal abilities rather than age, suggesting that their grammatical development does not follow the smooth relation to age that is found in TD children.

1. Background

Studies on the acquisition of relative clauses (RCs) have shown cross-linguistically that children struggle more with RCs headed by an object (1) than with those headed by a subject (2) (Sheldon, 1974; Roth, 1984; McDaniel, McKee & Bernstein, 1998; Berman, 1997; Friedmann & Novogrodsky, 2004; Adani et al., 2010, Arosio et al, 2010; Adani, 2011; Costa et al., 2011).

   (1) Show me the bear that the elephant is pushing __   (intervention/inclusion)

   (2) Show me the bear that __ is pushing the elephant   (no intervention)

Asymmetric performance for object versus subject RCs also surfaces in adults, in the form of both slower processing (Cook, 1975; Ford, 1983; Frauenfelder et al., 1980; Hakes et al., 1976) and avoidance in production (Contemori & Belletti, 2013; Costa et al., 2011).

    Grammatical approaches have accounted for the selective difficulty with object RCs in terms of locality constraints (Friedman et al., 2009; Contemori & Marinis, 2013). In an object RC such as (1), the fronting of the object 'the bear' entails crossing over a subject 'the elephant' sharing similar grammatical features, here +NP because both are full NPs. Put differently, the inclusion of the +NP feature on the intervening (i.e. c-commanding) lexical subject within the feature set of the target creates a dependency that is costly to parse for the computational system. The absence of an intervener, as in a subject RC (2), or a disjunction of features with the intervener, as in the 'free' object RC in (3) (where the moved element 'who' does not share the +NP feature with the subject 'the elephant'), renders the structure easier to parse:

3

(3) Show me who the elephant is pushing __   (intervention/disjunction)

The locality effects observed here strike an analogy with a principle that is operative in adult grammar known as Relativized Minimality (RM) (Rizzi, 1990; 2004, 2013; Starke, 2001). This principle constrains syntactic computations, along the lines of principles of Minimal Search (Chomsky, 2000, 2001). Essentially the RM principle states that a local relation between X and Y cannot hold if there is an intervener, Z, which is of the same structural type as X and which can be a potential candidate for the relation (4):

(4)   X   Z   Y

         Target   Intervener   Origin
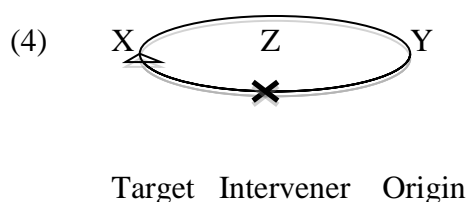
Intervention is defined in hierarchical terms through c-command: Z structurally intervenes between X and Y when Z c-commands Y and Z does not c-command X.

RM was initially developed to account for the ungrammaticality of structures like (5), which are different to (2) in that they are not simply costly to parse, but straightforwardly ruled out by the grammar.

(5) *What do you wonder who built __?

4

According to RM, movement of *what* in (5) is blocked by the structural intervention of *who*, a Wh-operator that is featurally identical to the head of the chain, i.e. both *who* and *what* are +WH operators, and the structure is thus ungrammatical. In (1), the head of the OR is not featurally identical to the intervening element because it is specified for both +R (being a relative operator) and +NP (being lexically restricted), while the intervening subject shares only the +NP specification. Still, these various phenomena of both child and adult grammars can be captured by recent developments of the featural RM system, where set-theoretic relations established involve different levels of complexity depending on the interaction between features: The relation of featural identity (as in 5') is straightforwardly ungrammatical, that of inclusion (as in 1') is grammatical but difficult, and finally that of disjunction (as in 3') is the easiest one.

(5')   *What do you wonder who built __?   (identity)
         Wh   **Wh**

(1')   Show me the bear that the elephant is pushing __   (inclusion)
              R **NP**   **NP**

(3')   Show me who   the elephant is pushing __   (disjunction)
             Wh R   NP

Recently, it has been observed that while a similarity in lexical NP restriction between the head and the intervener yields worse performance in children, a mismatch in features such as number and gender on the head of the dependency has

the potential to improve their comprehension, with effects emerging differently across languages. For example, a mismatch in gender between the lexically restricted head of a relative clause and the intervening subject does not significantly improve children's performance with these structures in Italian (Adani et al., 2010), however it does so in Hebrew, arguably because gender is expressed in the verbal inflection of Hebrew but not Italian (Belletti et al. 2012). The system of set-theoretic relations is thus enriched to include instances referred to as 'intersection', where the head and the target share some, but not all the relevant features. This is illustrated in (6) below, such that while the +NP feature below is shared, the value of number here is distinct (see Belletti et al. 2012: 1063):

(6) Show me the bear that   the elephants are pushing   (intersection)

           R NP Sg   NP Pl

It is important to point out that even if number were inaudible in the example above, the effect should nevertheless be present: overt inflection is not expected to be a necessary criterion for a feature to have an impact because inaudible features are also relevant for syntactic movement. However, this remains to be tested because the features studied so far, (gender in Hebrew, number in Italian,) have always had a corresponding morphological manifestation on the clausal inflectional head. It also remains to be determined to what extent sensitivity to these featural mismatches shows a developmental effect. Some authors report that already very young children are able to capitalize on distinctions in grammatical features leading to intersection relations (Belletti et al 2012). As a result, the child's system copes well once there is a difference in one relevant feature, leading to 'non-inclusion'. However other authors

report that performance with intersection relations shows a main effect of age (Adani 2010), possibly because sensitivity to grammatical mismatches in RCs relates to memory span, whose capacity also increases with age (Arosio et al. 2010; 2012). Very young children with limited computational resources should thus struggle to capitalize on featural mismatches more than older ones. Pursuing this reasoning, the comparison of different degrees of featural overlap, as in the case of intersection, is arguably an operation which is more complex than computing disjunction, so mastery of the former should be susceptible to developing later (see e.g. Friedman et al 2009: 84). The four set-theoretic relations are thus expected to be ranked in a precise gradient of distinctness as follows: identity < inclusion < intersection < disjunction.

In sum, structures involving intervention (as in object dependencies) have been shown to be more difficult than those without this configuration of intervention (subject dependencies), in particular when both the moved element and the intervening one have an NP feature, making them more similar in terms of their feature sets. Researchers working on RM have further observed that fine-grained featural differences of the intervener may improve parsing despite a similarity in NP, providing the distinct features are amongst those which are relevant for syntactic movement. Throughout, complexity is claimed to arise depending not on movement per se but on the similarity of the moved element and the intervener.

Movement in itself also plays a role in rendering a given structure complex and is expressed by the Derivational Complexity Metric (DCM) provided in (7) (Jakubowicz 2004, 2005, 2011).

(7) a. Merging **α**i n times gives rise to a less complex derivation than merging **α**i

(n + 1) times.

b. Internal Merge of **α** gives rise to a less complex derivation than Internal

Merge of **α** + **β**.

The central notion of the DCM is that Merge operations increase complexity,

including Internal Merge, i.e. movement. Research has empirically illustrated this

with respect to production data on the acquisition of wh-questions in French, where

children initially perform better with in-situ structures (8) than ex-situ ones (9)

(Hamann, 2006; Jakubowicz, 2004, 2005; Strik, 2008; van Kampen, 1997;

Zuckerman, 2001).

(8) Tu as poussé qui?

2p have pushed who

(9) Qui tu   as poussé?

who 2p have pushed

Children thus start by producing the least complex option, namely *wh* in-situ, before

moving on to more computationally complex structures involving *wh*-fronting:

Athough there are reports of individual variation, French-speaking children tend to

commence producing *wh*-questions around the age of 2;0 (Hulk 1996; Plunkett 1999;

Hamann 2006), and around the age of 4 to 5 years the preference is reversed, giving

rise to more questions with *wh* ex-situ being produced than with *wh* in-situ (Hulk &

Zuckermann 2000).

Few studies have examined whether the complexity added by movement

8

reported for production carries over to comprehension. Moreover an assessment of the contributory roles played by intervention, on the one hand, and movement, on the other, is also largely missing from the literature. Related to this, it remains to be seen how children fare with the parsing of in-situ *which*-NP questions, such as (10), as compared to in-situ *who*-NP questions (11), and ex-situ *which/who*-NP questions (12,13), all structural options in French.

(10)   Les éléphants poussent quel ours?

         The elephants push which bear?

(11)   Les éléphants poussent qui?

         The elephants push who

(12)   Quel ours les éléphants poussent?

         Which bear the elephants push?

(13)   Qui les éléphants poussent?

         Who the elephants push?

Exploring the comprehension of both ex-situ and in-situ questions can help to further elucidate the scale of complexity impacting parsing in acquisition, and may also shed light on the syntactic analysis assigned to wh in-situ. Indeed there are diverging views in the literature on this topic. On the one hand, structures such as (10,11) above have been argued to have the same logical form as that of their ex-situ counterparts (Huang, 1982), i.e. 'for what x [... x ...]'. This view implies that wh in-situ elements undergo covert phrasal movement and are thus on a par with quantifiers, ultimately producing the relevant operator-variable structure. If so, one would expect children to struggle with intervention effects not only with ex-situ questions such as (12) but also

with in-situ questions such as (10). Other approaches suggest that *wh* in-situ involves movement of the *wh*-feature only (Mathieu 1999; Cheng and Rooryck 2000, Baunaz 2011; Shlonsky 2012)[1]. If so, in-situ structures would be better processed because avoiding pied-piping of the entire phrase would in turn circumvent intervention effects. Acquisition of these structures in French offers a promising testing ground for these diverging theoretical perspectives. In addition, given the recent evidence of intervention effects mentioned above, it is necessary to assess the role of movement independently of intervention in order to evaluate the relative role of these two factors independently from each other in a single experiment with the same participants. Finally, it is important to compare the acquisition of wh-questions with that of other dependencies involving movement and intervention. The analysis of complexity explored here remains the same not only for questions but also for relative clauses (RCs, see examples 13 and 14 for an illustration from French). Indeed RCs involving movement and intervention (14) are reportedly acquired later than those with only movement but not intervention (13) (Friedmann et al., 2009; Adani 2011):

(14) Les éléphants qui poussent les ours

The elephants who push the bears

(15) Les ours que les éléphants poussent

The bears who the elephants push

Example (13) illustrates a subject relative clause, with the relative pronoun *qui*

---

[1] See also Reinhardt (1997) for an earlier analysis avoiding fronting of a *wh*-operator thanks to insertion of a question operator in the base position which binds a choice function.

designating the subject of the clause (*les elephants:* the elephants). Example (14) is an

object relative clause (OR), with the pronoun *que* designating the object of the clause

(*les ours:* the bears). Early reports of the acquisition of these structures in French

suggested the frequent use of resumptive pronouns in RCs as compared to Wh-

questions, and analysed this as indicative of the later emergence of movement in RCs

as compared to Wh-questions (Labelle 1990, 1996), although this analysis has not

been adopted in more recent studies (Guasti & Shlonsky 1995; Guasti & Cardinaletti

2003). Like for *wh*-questions, studies assessing the development of RCs specifically

in French have mainly focused on production. The current work contributes to

bridging the gap in studies of A'-dependencies by assessing comprehension of both

RCs and *Wh*-questions in French, via the use of a specific complexity metric.

Studying participants with typically developing language can enlighten us on

how these syntactic factors emerge in a typical population, and can further provide a

benchmark with which to assess atypically developing populations, so as to determine

whether or not their course of development is merely delayed but does not differ

fundamentally from unimpaired populations. One population for whom there is a

debate on these issues is Autism Spectrum Disorder (ASD). Grammatical

development is under investigated in ASD, and the reports thus far have yielded

conflicting findings. Some authors have argued that the acquisition of syntax is

delayed but globally reminiscent of that attested in TD children matched on cognitive

abilities (Tager-Flusberg et al., 1990; Lord & Paul, 1997). Other authors have claimed

that grammatical delays, in particular with non-canonical structures, may be unrelated

to the children's general cognitive abilities (for object relative clauses see Riches et

al., 2010, for object questions see Zebib et al., 2013, for object clitics see Terzi et al.,

2014). Certain reports agree that patterns of grammatical development in this population are difficult to account for in terms of general cognitive difficulties, however reveal that difficulties emerge also with canonical structures, and even with structures which are unproblematic at any stage of typical acquisition, hence suggesting a deviant developmental path (see e.g. Perovic et al., 2007; Perovic et al., 2013 for difficulties in ASD with constraints on binding and reflexive pronominal interpretation). More work is thus necessary to gain insight into the formal language profiles of children with ASD, including with respect to how they deal with potential sources of syntactic complexity such as intervention and movement. To address whether difficulties in grammatical development are due to general cognitive difficulties, it is crucial to compare the performance of children with ASD to control children who are matched on their non-verbal abilities.

2. Aims of the present study

The present study has two aims. The first aim is to investigate the effect of syntactic complexity in the development of A'-dependencies in French-speaking children of a wide age-range based on the metric articulated in Table 1. The metric involves three cumulative factors defining four increasing levels of complexity[2].

Add Table 1 around here

---

[2] The design here considers movement and different types of (potential) interveners but does not include other factors, which have been claimed to increase complexity, such as embedding (see Owen and Leonard 2006, Delage et al. 2008, Friedman, Yakini and Szterman in print). We have reasons to doubt the impact of embedding per se. Indeed our data do not support this, as we will show in section 4.

The test cases involve questions and relative clause structures spread across the 4 levels defined by the independent variables: movement, configurational intervention from an element which is featurally different from the moved element (the features of the moved element are disjoint from the features of the intervener) and configurational intervention from an element which is featurally similar to the moved element (the features of the moved element intersect with those of the intervener; see Methods section for details on the design). If the complexity metric reflects principles of the grammar, we predict that accuracy should significantly decrease as the number of complexity factors increases. Complexity may affect children differently depending on their developmental stage. We hypothesize that the three factors involved in the complexity metric (movement, intervention, and similarity) will affect the children's performance, but it may not affect their performance in an equal way..

The second aim addresses the impact of the complexity metric and the three factors in children with ASD. The metric in Table 1 serves as a tool for assessing whether syntactic development in children with ASD follows the same pattern as that of younger TD children who have similar non-verbal abilities, and whether their performance is related to their age and/or their non-verbal abilities. If children with ASD have a delay in the acquisition of syntax, the three factors should affect the children with ASD in a similar way as the younger TD control children. In contrast, if children with ASD follow a different developmental trajectory than TD children, the three factors may affect the children with ASD differently than the younger TD control children.

13

3. Methodology

Participants

Our participants included forty-five TD native French-speaking children from three school levels (mean ages 4;9, 6;8 and 8;8), as shown in Table 2. This enabled us to investigate how movement and intervention develop over time by comparing the performance of the three age groups, and to assess if some factors play a more crucial role at one or another stage of development[3]. The TD children were recruited from a primary school in Geneva, Switzerland.

Add Table 2 around here

In addition, 15 children with ASD aged 6 to 16 (mean age 9;4) also participated.[4] The children with ASD were recruited through parent associations and psychologists in the Geneva region and had been previously diagnosed by a specialist as meeting DSM-IV criteria for an autism spectrum disorder (American Psychiatric Association 1994). A subgroup of 15 TD children from the three groups mentioned above was selected to be matched with the children with ASD on non-verbal reasoning (Raven's Progressive Matrices; Raven et al. 1986) [$F_{(1, 29)} = 0.236$, $p = 0.631$, $\eta p2 = 0.009$]. Matching the two groups on their non-verbal abilities is very important in research on

---

[3] See also Adani, et al. (2010) who explore related issues in a sample of comparable age range, which furthermore allows cross-linguistic comparisons.

[4] Given that previous studies investigating the language abilities of children with ASD found difficulties in structures that are early acquired (e.g. Perovic, et al., 2013; Terzi, et al., 2014) we did not include children with ASD younger than 6 years old to avoid a floor effect in the structures with high complexity. The wide age range enabled us to include children with a wide range in their performance in the experimental task.

14

ASD (see Tager-Flusberg 2004). If children with ASD have lower non-verbal abilities than TD children, a difference between the two groups in their verbal abilities can be the result of differences in their non-verbal abilities. By matching the two groups for their general non-verbal abilities, we will be able to address differences between them that cannot be attributed to differences in their non-verbal abilities. The TD children were younger than the children with ASD [$F_{(1, 29)} = 4.249$, $p = 0.049$, $\eta p2 = 0.136$]. Table 3 shows the children's age and scores on the Raven's.

Add Table 3 around here

Testing for the TD group was conducted in a quiet classroom of their school and the group with ASD was tested in the comfort of their own homes. The experiment and cognitive testing was run in two sessions of roughly 30 minutes and children were offered a break in the middle. Approval for this study was obtained from the Ethics Committee of the Psychology Department of the University of Geneva and parents of participants provided informed, written consent for their children to participate in the study.

Material

The participants took part in a sentence-picture matching task. Table 4 presents examples of the test sentences for each sentence type and condition based on the complexity matrix presented in Table 1.

Add Table 4 around here

Level 0 consisted of two sentence types that do not have movement, namely Object questions in situ without an NP object and Object questions in situ with an NP object. Level 1 consisted of three sentence types that have movement but no intervention, namely Subject questions without an NP subject, Subject questions with an NP subject, and Subject relatives with an NP subject. Level 2 consisted of Object questions ex situ that have movement and intervention from a dissimilar element, i.e., the featural make-up of the moved object (+WH/+R, -NP, +Sg) and that of the intervening subject (+NP, +Pl) are fully disjoint. Finally, Level 3 consisted of object questions ex situ and object relatives with an NP object and an NP subject, two sentence types that include movement, intervention, and similarity, i.e., the featural make-up of the moved object ((+WH/+R, +NP, +Sg) and that of the intervening subject (+NP, +Pl) intersect.[5]

The characters referred to by the wh-question or RC were singular or plural, and this was counterbalanced across conditions. The difference between singular and plural in the verbal inflection was always inaudible for the verbs used. All pictures (taken from Adani 2011) followed the same basic format, with (an) animal(s) X on the left, (an) animal(s) Y in the middle and (an) animal(s) Z on the right. External characters were always of the same type, e.g., a bear that is pushing two elephants and

---

[5] Note that the only sentence types that include embedding in our design are subject and object relative clauses. If embedding affects the children's performance, this should lead to lower performance in these particular sentence types compared to the sentence types within the same level.

these two elephants are pushing another bear (see Figure 1).

Add Figure 1 around here

The image in Figure 1 was paired with one of the structures in Table 4, i.e. either a RC or a *wh*-question, pronounced online by the experimenter (see Appendix for detailed protocols). The correct answers were always located on one of the peripheries of any given picture. All items were semantically reversible, so that the child could not simply rely on understanding of the words alone to understand the sentence.

Procedure

The sentence-picture matching task was computerized and presented on a 15" portable computer, with the keyboard covered so as to avoid distracting participants. The total testing involved two experiments, one containing 24 RCs and the other 48 *wh*-questions, all sentences having been pre-recorded and administered using headsets connected to the laptop. Each experiment was divided into two equal parts to allow for a break in the middle of the testing and thus keep participants engaged. The RCs were divided into 12 object relatives and 12 subject relatives with a lexically-restricted (+NP) head. These were preceded by 4 familiarization items ensuring that participants knew what they had to do, and interspersed with 16 simple imperatives used as fillers, ensuring children were not overloaded and that they felt successful (e.g. "Show me the girl with the bow"). There were also 48 *wh*-questions, divided between +NP and –NP (i.e. bare) *wh*-elements, and in-situ versus ex-situ structures.

These were preceded by 6 familiarization items and interspersed with 24 simple imperatives that served as fillers. At the very beginning of the testing, there was a lexical training phase including corrective feedback so as to check that children knew the lexicon necessary to take part in the experiment. It was common to both the RC and *wh-* experiments and consequently only administered once. The ordering of the tasks was counterbalanced across children and items were randomized via e-prime throughout the two experiments.

The experiment began with the following oral instructions from the experimenter: "You will see pictures on the screen, and a voice will ask you to show one of the characters in the picture. Listen carefully and look at all of the characters, then point to the answer you think is the best." During the testing phase, no feedback was given when the child pointed to the wrong character except for commending them for their efforts (e.g. "Good listening!"). Children were presented with one picture at a time on the computer screen and heard one sentence. They were left as much time as they needed to answer and the experimenter ensured that she had their full attention before moving on to the following stimulus. In some cases, the children were asked to press a button on the computer to move on to the following stimulus, as this helped them to keep concentrated on the task at hand.

Scoring and data analysis

The criteria for excluding children from the data analysis were failure to learn the characters during the lexical phase and/or below or at-chance performance in the filler items. All children succeeded in the lexical phase apart from one TD child who was excluded due to poor results on the filler items. The dependent variable was accuracy

rate, as measured by whether or not the character that the child pointed to after having heard the target sentence was correct.

Descriptive statistics were conducted first per structure and complexity level to explore the data. These were followed by inferential statistics using mixed repeated measures ANOVAs with Group as the between factor and Complexity Level as the within factor to investigate differences between the groups and the levels of complexity and address the research questions. Interactions between Group and Complexity Level were followed up using within group analyses for each group separately and between group analyses for each level of complexity. Follow up analyses were conducted within ANOVAs using Bonferroni correction. All ANOVAs included partial Eta squared analyses (ηp^2) that provide information about the effect size.[6] To investigate whether syntactic complexity in children with ASD develops as a function of age and/or non-verbal abilities, Pearson correlations were conducted between each level of complexity, age and non-verbal abilities. Separate analyses were conducted to address the four aims of the study. The first set of analyses compared the three groups of TD children to address the development of the four levels of complexity. The second set of analyses addressed effects of movement intervention, and feature (NP) similarity separately within TD children. For effects of movement, we compared the two structures that did not involve movement (Level 0) to all other structures that involved movement (Levels 1-3). For effects of intervention, we compared the structures with movement but without intervention (Level 1) with the structures with movement and intervention (Levels 2 and 3). For effects of feature similarity, we compared the structures with intervention and a

---

[6] Partial Eta square demonstrates the proportion of the variance that is attributable to the factor tested.

disjunction of features (Level 2) with the structures with intervention and an intersection of features (Level 3).

The third set of analyses compared the children with ASD to their matched controls to investigate whether computational complexity impacts children with ASD in a similar way as TD children who have similar non-verbal reasoning abilities. The final set of analyses compared the children with ASD to the TD controls for effects of movement, intervention, and feature similarity.

4. Results

The effect of complexity in the development of A'-dependencies in TD children

Table 5 shows the descriptive statistics for accuracy in the comprehension of the three groups of 4-, 6-, and 8-year old TD children per structure and Figure 2 illustrates their accuracy per level of complexity.

Add Table 5 around here

Add Figure 2 around here

The descriptive statistics in Table 5 show overall a large degree of consistency of the structures within each level of complexity. The repeated measures ANOVA revealed a significant main effect of Group [$F_{(2, 42)} = 27.952$, $p < 0.001$, $\eta p2 = 0.571$], a significant main effect of Complexity Level [$F_{(3, 126)} = 36.598$, $p < 0.001$, $\eta p2 =$

0.466], and a significant interaction between Group and Complexity Level [F (6, 126) = 5.452, p < 0.001, ηp2 = 0.206], indicating that the three groups of TD children did not show the same pattern of performance in the four levels of complexity.

The within group analyses showed effects of syntactic complexity in all three groups of TD children. All groups had lower performance in sentences from Level 3 compared to Level 0 [4-year olds: p < 0.001; 6 year olds: p = 0.001; 8-year olds: p < 0.001]. However, 4- and 6-year olds showed also significant differences between Level 0 and Level 2 [4-year olds: p = 0.022; 6 year olds: p = 0.002], Level 1 and Level 2 [4-year olds: p = 0.04; 6 year olds: p = 0.003], and Level 1 and Level 3 [4-year olds: p < 0.001; 6 year olds: p = 0.004].

The between group analyses showed clear developmental differences between the age groups. This was evident at Level 0 [4- vs. 6-year olds: p < 0.001; 4- vs. 8-year olds: p < 0.001], Level 1 [4- vs. 6-year olds: p = 0.002; 4- vs. 8-year olds: p = 0.001], and Level 3 [4- vs. 6-year olds: p < 0.001; 4- vs. 8-year olds: p < 0.001]. At Level 2, 4-year olds performed similarly to 6-year olds, but less well than 8-year olds [4- vs. 8-year olds: p = 0.013]. 6-year olds performed as well as 8-year olds at Levels 0, 1, and 3, but not at Level 2 [p = 0.013].

The development of movement, intervention, and feature similarity in TD children

Effect of movement. Figure 3 compares the accuracy in the comprehension of structures without movement compared to structures involving movement.

Add Figures 3-5 around here

21

The repeated measures ANOVA revealed a significant main effect of Group [F (2, 42) = 25.731, p < 0.001, ηp2 = 0.551] reflecting lower accuracy in 4- than 6- [p < 0.001] and 8-year olds [p < 0.001], but no significant difference between 6- and 8-year olds. There was also a significant main effect of Movement [F (1, 42) = 47.307, p < 0.001, ηp2 = 0.530] indicating lower performance in structures with movement compared to the ones without.

Effect of intervention. Figure 4 compares the accuracy in the comprehension of structures with movement but without intervention compared to structures with movement and intervention . The repeated measures ANOVA revealed a significant main effect of Group [F (2, 42) = 33.754, p < 0.001, ηp2 = 0.616], a significant main effect of Intervention [F (1, 42) = 68.721, p < 0.001, ηp2 = 0.621] and a significant interaction between Group and Intervention [F (2, 42) = 8.254, p = 0.001, ηp2 = 0.282], indicating that the three groups were not affected equally by intervention effects. Within subjects analyses showed intervention effects in all three groups, but the effect was stronger in 4- [F (1, 14) = 33.592, p < 0.001, ηp2 = 0.706] and 6-year olds [F (1, 14) = 32.034, p < 0.001, ηp2 = 0.696] than in 8-year olds [F (1, 14) = 7.417, p = 0.016, ηp2 = 0.346]. Between subjects analyses showed differences between 4- and 6-year olds [p = 0.002] and 4- and 8-year olds [p = 0.001] in structures without intervention [F (2, 44) = 10.329, p < 0.001, ηp2 = 0.330], and with intervention [F (2, 44) = 22.700, p < 0.001, ηp2 = 0.519; 4- vs. 6-year olds: p < 0.001; 4- vs. 8-year olds: p < 0.001], but the effects in structures with intervention were stronger than the ones without and no significant difference was attested between 6- and 8-year olds.

22

Effect of feature similarity. Figure 5 compares the accuracy in the comprehension of structures with intervention and disjoint feature sets compared to structures with intervention and intersecting feature sets. The repeated measures ANOVA revealed a significant main effect of Group [$F_{(2, 42)} = 17.473$, $p < 0.001$, $\eta p2 = 0.454$], and a significant interaction between Group and Feature Similarity [$F_{(2, 42)} = 5.710$, $p = 0.006$, $\eta p2 = 0.214$], indicating that the three groups were not affected equally by feature similarity in structures with intervention. Within subjects analyses showed that there was an effect of feature similarity only in 4-year olds [$F_{(1, 14)} = 7.356$, $p = 0.017$, $\eta p2 = 0.344$]. Between subjects analyses showed that in structures with feature dissimilarity, 8-year olds performed better than 4- [$p = 0.013$] and 6-year olds [$p = 0.047$] [$F_{(2, 44)} = 5.250$, $p < 0.009$; $\eta p2 = 0.200$], but there was no difference between 4- and 6-year olds. In contrast, structures with feature similarity showed that 4-year olds performed less well than 6- [$p < 0.001$] and 8-year olds [$p < 0.001$] [$F_{(2, 44)} = 25.437$, $p = 0.009$; $\eta p2 = 0.548$], but there was no difference between 6- and 8-year olds.

The effect of complexity in the development of A'-dependencies in children with ASD compared to TD controls

Table 6 shows the accuracy in the children with ASD compared to their TD controls per structure and Figure 6 illustrates their accuracy per level of complexity.

Add Table 6 around here

23

Add Figure 6 around here

Table 6 shows overall a large degree of consistency of the structures within each level of complexity. The repeated measures ANOVA revealed a significant main effect of Group [$F_{(1, 27)} = 6.664$, $p < 0.016$, $\eta p2 = 0.198$] and a significant main effect of Complexity Level [$F_{(3, 81)} = 15.011$, $p < 0.001$, $\eta p2 = 0.357$]. There was no significant interaction between Group and Complexity Level. This indicates that the children with ASD follow the same pattern of performance in the four levels of complexity with the TD children.

The within group analyses showed a similar pattern to the performance of 4- and 6-year olds in the previous analyses. There was a lower performance in sentences from Level 3 compared to Level 0 [$p < 0.001$] and Level 1 [$p = 0.001$] and also in sentences from Level 2 compared to Level 0 [$p = 0.001$] and Level 1 [$p = 0.018$]. Moreover, there was no significant difference between sentences at Level 0 vs. 1 and also Level 2 vs. Level 3.

To investigate if the children's performance on the comprehension task relates to their age and non-verbal abilities, Pearson's correlations were conducted between the children's age, non-verbal IQ and the four levels of complexity for each group separately. The results are illustrated in Table 7.

Add Table 7 around here

24

In the group of TD children, non-verbal IQ correlated with age and all correlations between age/non-verbal abilities and levels of complexity (apart from Level 2) were significant. In the children with ASD, on the other hand, non-verbal IQ did not correlate with age. Non-verbal IQ correlated with all levels of complexity (for Level 0 it was approaching significance, $p = 0.051$), but age correlated only with Level 2. All correlations were positive, indicating that older children and/or children with higher non-verbal abilities had better accuracy than younger children and/or children with lower non-verbal abilities.

The effect of movement, intervention, and feature similarity in children with ASD compared to TD controls

Effect of movement. Figure 7 compares the accuracy in the comprehension of structures without movement compared to structures involving movement.

Add Figures 7-9 around here

The repeated measures ANOVA revealed a significant main effect of Group [$F (1, 27)$ = 6.434, $p = 0.017$, $\eta p2 = 0.192$] reflecting lower accuracy in children with ASD compared to TD controls. There was also a significant main effect of Movement [$F (1, 27) = 34.830$, $p < 0.001$, $\eta p2 = 0.563$] indicating lower performance in structures with movement compared to the ones without, and no significant Group by Movement interaction.

25

Effect of intervention. Figure 8 compares the accuracy in the comprehension of structures with movement but without intervention compared to structures with movement and intervention. The repeated measures ANOVA revealed a significant main effect of Group [$F_{(1, 27)} = 8.419$, $p = 0.007$, $\eta p2 = 0.238$], reflecting lower accuracy in children with ASD compared to TD controls. There was also a significant main effect of Intervention [$F_{(1, 27)} = 22913$, $p < 0.001$, $\eta p2 = 0.459$] indicating lower performance in structures with intervention compared to the ones without, and no significant Group by Intervention interaction.

Effect of feature similarity. Figure 9 compares the accuracy in the comprehension of structures with intervention and dissimilar features (disjunction) compared to structures with intervention with similar features (intersection). The repeated measures ANOVA revealed no significant main effects or interactions, indicating similar performance in the two groups and structures.

Discussion

The study had two aims: 1) to investigate the effect of complexity in the development of A'-dependencies in French-speaking children by addressing how each one of the three factors involved in the metric (movement, intervention, and feature similarity) impacts the development of A'-dependencies, 2) to investigate the impact of the complexity metric and the three factors in children with ASD compared to TD controls[7].

---

[7] Recall in footnotes 1 and 4 we explained that our design does not test embedding as a measure of complexity, and therefore, we included embedding only in two out of the eight conditions. Indeed Friedmann et al (2009) in their study of relative clauses and questions in Hebrew report better performance for the headed subject relative clauses (92%) than which subject questions (75%), which is the opposite pattern we would have expected if embedding negatively influenced the comprehension of these structures. Our ASD population also shows a tendency towards more difficulty with headed

26

The complexity metric was based not only on movement (expressed by the DCM), but also on the nature of intervention and the feature similarity between the moved element and the intervener (based on RM). A range of eight structures were included in this study that involve structures with and without movement; from the structures that involve movement, structures with and without intervention; and finally from the structures that have movement and intervention, structures with and without feature similarity between the element that moves and the intervener. The eight structures were classified on the basis of these three cumulative factors in four levels of increasing complexity.

The effect of complexity in TD children

The results from the TD children demonstrate that comprehension is overall affected by syntactic complexity operationalised in terms of Movement, Intervention, and Featural similarity and a considerable degree of consistency for the structures within each level of complexity. However, the pattern of performance differed as a function of the children's age. All groups had lower performance in the most complex structures (Level 3) compared to the simplest ones (Level 0), however, age groups differed in their sensitivity to the intermediate levels. The 8-year old children had high accuracy so that intermediate levels that did not differ from each other. In contrast, 4- and 6-year old children had low performance also in structures at Level 2 that differed from Levels 0 and 1. Movement affected all three groups in a similar way, but intervention effects were strongest in 4-year olds and became weaker as a

wh-questions (77.2%) than with headed RCs (82.8%) (so, once again the opposite pattern to what would be expected if embedding was the key measure of complexity). As for our TD sample, headed subject RCs and headed subject questions are equivalently well-understood (95.2% for subject questions and 96.4% for subject RCs).

function of age. Feature similarity affected only 4-year olds.

These findings support the hypothesis that movement, intervention and feature similarity all play a significant, cumulative role in children's sentence comprehension, although to various degrees depending on children's linguistic development. That intervention effects arise even in the case of a featural disjunction between the moved element and the intervener suggests that configurational intervention has a role to play independently of featural similarity, even though similarity increases its effect, proportionally to the feature overlap: the more they overlap the more the intervener perturbs the dependency[8].

Another finding of the present study was that structures with two NPs and therefore intersecting feature sets were not harder to process than those with disjoint feature sets for older children. The fact that having to full NPs failed to increase difficulty in that group may be due to their ability to exploit morphosyntactic differences in terms of number information on the NPs, in line with what has been reported for Italian (Adani et al., 2010, 2013; Arosio et al., 2012; see Franck et al., 2015 for a similar finding in French speaking adults). Interestingly, while number is systematically audible on the verbal inflectional head in Italian, it was never the case

---

[8] Featural disjunction has been used to describe structures such as ours where an operator without lexical restriction (e.g. who) crosses over a lexically restricted subject (Friedmann et al., 2009). Indeed in these structures, there is no similarity for syntactic features which are part of the featural composition of the clausal inflectional head, a criterion claimed to render a feature relevant for the computation of locality (Belletti et al. 2012: 1062). However another view is put forth in Costa et al. (2012). These authors also report effects of intervention in the absence of lexical restriction on the head of the chain, and claim that there is nevertheless featural similarity with the intervener, e.g. for animacy. This would give rise to attenuated (but detectible) effects compared to those that arise with a similarity in lexical restriction. Although animacy is not part of the feature set of the clausal inflectional head, this feature has been argued to be relevant for movement (see Bianchi 2006), which may then be the crucial ingredient for a feature to enter into the computation of locality, rather than a feature being part of the featural composition of the clausal inflectional head. For more on this perspective, see Bentea, Durrleman & Rizzi (2016).

on the French verbs included in the present study. That number mismatch can facilitate parsing, independently of the morphological realization of the feature follows from a feature-based account, where inaudible features remain relevant for movement and thus for locality.

The effect of NP feature similarity in 4-year-old children demonstrates that similarity in terms of lexical restriction negatively affects children's comprehension at an early stage of development, and that contrarily to older children, 4-year-olds cannot reliably capitalize on the number for the computation of locality. This finding is compatible with findings by Adani et al., (2010, 2013) who tested children of 5 and over when showing that number mismatch improved performance with A'-dependencies, and it is also compatible with reports that sensitivity to morphological number cues emerge only after age 5 (Johnson et al. 2005; Pérez-Leroux 2005; Miller and Schmitt 2009)[9]. Along these lines, it has been argued that the sensitivity to such grammatical features may related to memory span, whose capacity increases with age (Arosio et al. 2010; 2012).

The children's accuracy for *wh* in-situ questions regardless of the lexical restriction on the *wh*-head provides support for an analysis of these structures in terms of featural movement (Mathieu, 1999; Cheng & Rooryck, 2000; Baunaz, 2011; Shlonsky, 2012) rather than covert phrasal movement involving pied-piping of the entire object DP (along the lines of Huang, 1982). If covert movement were equivalent to overt movement, one would predict the same intervention effects to

---

[9] Belletti et al (2012) studied Hebrew participants aged 3;9- 5;5 years who were reported to fare well with structures sharing a lexical NP feature yet differing in gender (81% accuracy), which may seem at odds with our findings (70% accuracy). However this discrepancy could be due to a difference in tasks: their task involved the selection of an image amongst two, images depicting participants carrying out actions in reversed thematic roles, thus yielding a 50% chance of success. Our material involved the selection of the correct character amongst three possibilities, thus there was a 33% chance of success.

arise for both object questions involving *wh* in-situ and those involving *wh* ex-situ. This prediction was not borne out by our results. As such, an analysis in terms of movement of the [+WH] feature on the *wh*-element in-situ, rather than of the whole set of features present on this object, has the potential to account for the absence of intervention effects reported here for *wh* in-situ questions such as (16):

(16) L'ours pousse   quel éléphant?

   The bear pushes   which elephant

   [+NP]   [+**WH** +NP]

The effect of complexity in children with ASD

The data from the present study provide important information on the syntactic profile of children ASD, which has been only seldom explored to date. Previous studies on the grammatical abilities of children with ASD have demonstrated subtle difficulties in the grammatical abilities of children with ASD compared to TD control children (Perovic et al., 2007, Perovic et al., 2013; Riches et al., 2010; Terzi et al., 2014; Zebib et al., 2013), but the groups of children with ASD and TD controls were not always matched in the same way, with some studies matching the groups on age and verbal abilities, but not on general cognitive abilities (Terzi et al., 2014), others matching them on age but not on verbal and non-verbal abilities (Riches et al., 2010), and others matching them on non-verbal and verbal abilities but not on age (Perovic et al., 2013). To address whether syntactic development in children with ASD follows the same pattern as that of TD children who have similar non-verbal abilities and whether their performance is related to their age and/or their non-verbal abilities, we compared the performance of children with ASD to a group of younger TD children who were

matched on their non-verbal abilities. The results showed that the children with ASD had overall lower performance than their younger TD controls who had similar non-verbal abilities. This indicates that non-verbal abilities per se cannot account for the difficulties children with ASD have in grammar because if this were the case, the children with ASD would not have differed from their TD controls matched on non-verbal abilities. Furthermore, the children with ASD differed from the TD controls across all levels of complexity and not only in the most complex structures. This suggests difficulties across the board even for simple structures, which their younger TD peers have mastered and is consistent with studies matching children with ASD to TD children on non-verbal abilities and demonstrating that children with ASD have subtle difficulties also with simple linguistic structures which have not been reported to be problematic even in early stages of TD (Perovic et al., 2007; 2013).

Importantly, the performance of children with ASD was affected by syntactic complexity operationalised in terms of Movement, Intervention, and Feature similarity. More complex structures involving movement, intervention and similarity (Level 3) and structures involving movement and intervention with dissimilar featural sets (Level 2) were less accurate than less complex structures without movement, intervention, and feature similarity (Level 0) and with movement but without intervention and similarity (Level 1). This indicates that the syntactic complexity metric based on movement, intervention, and feature similarity plays a significant role not only in TD children, but also in children with ASD. The separate analyses addressing each one of these factors separately provided additional evidence for the way these factors affect the performance of the children with ASD compared to the TD control children. Movement and intervention affected both groups in a similar way and the children with ASD were less accurate than the TD controls across the

board, i.e., they were not affected disproportionally more than the TD children in conditions involving movement or intervention compared to the ones that did not involve movement and intervention. However, the ASD and the TD-matched children did not differ in accuracy and failed to show effects of feature similarity. This resembles the results of the older TD children in this study that did not show effects of feature similarity and suggests that they can exploit the number differences between the head and the intervener; number mismatches between the head and the intervener can facilitate parsing (see, Adani et al., 2010; 2013). The findings related to the complexity metric provide further support that the syntactic complexity metric adopted in the present study has psychological reality in both TD children and children with ASD.

The final important finding in the group of children with ASD regards the relationship between the children's age and non-verbal abilities and the scores on the syntactic complexity metric. Whereas the TD control children's performance developed as a function of the children's age and non-verbal abilities, which were related to each other, the ASD children's performance developed as a function of the children's non-verbal abilities and not age, given the absence of correlation with age. The correlation between non-verbal abilities and the ASD children's scores on the syntactic complexity metric indicates the importance of non-verbal abilities for the development of syntactic complexity. The lack of a correlation between age and the ASD children's scores on the syntactic complexity metric coupled with the lack of a correlation between age and non-verbal abilities suggest an asynchrony between age on the one hand and non-verbal abilities along with syntactic complexity on the other hand. Syntactic complexity in children with ASD does not seem to follow the smooth relation to age that is found in TD children.

References

Adani, F., van der Lely, H.K.J., Forgiarini, M., Guasti, M.T. (2010). Grammatical feature dissimilarities make RCs easier: a comprehension study with Italian children. *Lingua 120*, 2148–2166.

Adani, F. (2011). Rethinking the acquisition of RCs in Italian: towards a grammatically based account. *Journal of Child Language 38*, 141−165.

Arosio, F., Guasti, M-T., and Stucchi, N. (2010). Disambiguating Information and Memory Resources in Children's Processing of Italian Relative Clauses, *Journal of Psycholinguistic Research 40(2),* 137-54.

Arosio et al. (2012): Morphological Information and Memory Resources in children's processing of relative clauses in German. *Language Learning and Development 8*, 340-364.

Bianchi, V., (2006). On the syntax of personal arguments. *Lingua 116*, 2023--2067.

Baunaz, Lena. (2011). *The grammar of French quantification*. Dordrecht: Springer.

Bentea, A. Durrleman, S., Rizzi, L. (2016). Refining intervention: The acquisition of featural relations in object A-bar dependencies. *Lingua Vol 169*, 21-41.

Belletti, A., Friedmann, N., Brunato, D., Rizzi, L. (2012). Does gender make a difference? Comparing the effect of gender on children's comprehension of relative clauses in Hebrew and Italian. *Lingua*, 122, 1053--1069.

Berman, R. (1997). Early acquisition of syntax and discourse in Hebrew. In Y. Shimron (ed.), *Psycholinguistic studies in Israel : language acquisition, reading and writing*. Jerusalem: Magnes Press. (in Hebrew.)

Cheng, Lisa Lai-Shen and Rooryck, Johan. (2000). Licensing WH-in-situ. *Syntax 3,* 1-19.

Chomsky, N. (2000). Minimalist inquiries. In R. Martin, D. Michaels, & J. Uriagereka (Eds.), *Step by step*. Cambridge,MA:MIT Press.

Chomsky, N. (2001). Derivation by phase. In M. Kenstowicz (Ed.), Ken Hale: *A life in language*. Cambridge, MA: MIT Press.

Contemori, C., & Belletti, A. (2013). Relatives and Passive Object Relatives in Italian speaking children and adults: Intervention in production and comprehension. *Applied Psycholinguistics 36(6),* 1021-1053.

Contemori, C. & Marinis, T. (2013). The impact of number mismatch and passives on the real-time processing of relative clauses. *Journal of Child Language 17*, 1-32.

Cook, V.J., (1975). Strategies in the comprehension of relative clauses. *Language and Speech 18*, 204–212.

Costa, J., Lobo, M., Silva, C. (2011). Subject–object asymmetries in the acquisition of Portuguese relative clauses: Adults vs. children. *Lingua 121*, 1083–1100.

Ford, M., (1983). A method for obtaining measures of local parsing complexity throughout sentences. *Journal of Verbal Learning and Verbal Behavior 22*, 203–218.

Franck, J., Colonna, S., & Rizzi, L. (2015). Task-dependency and structure-dependency in number interference effects in sentence comprehension. *Frontiers in Psychology*. 6:349

Frauenfelder, U., Segui, J., Mehler, J., (1980). Monitoring around the relative clause. Journal of Verbal Learning and Verbal Behavior 19, 328–337.

Friedmann, N., Belletti, A., Rizzi, L. (2009). Relativized Relatives: types of intervention in the acquisition of A'-dependencies. *Lingua 119*, 67-88.

Friedmann, N., and Novogrodsky, R. (2004). The acquisition of RC comprehension in

Hebrew: A study of SLI and normal development. *Journal of Child Language 31*, 661-681.

Friedman et al. (2015) 'Relatively easy relatives: Children with syntactic SLI avoid intervention', in C. Hamann, E. Di Domenico and S. Matteini (eds)): *Structures, Strategies and Beyond: Studies in Honour of Adriana Belletti*. John Benjamins.

Guasti, M-T. and Cardinaletti, A. (2003). "Relative clause formation in Romance child's production." *Probus* 15: 47-89.

Guasti, Maria Teresa and Shlonsky, Ur. (1995). "The Acquisition of French Relative Clauses Reconsidered." *Language Acquisition 4:* 257-276.

Hamann, C. (2006). Speculations about early syntax: The production of wh-questions by normally developing French children and French children with SLI. *The Catalan Journal of Linguistics 5,* 143-189.

Hakes, B., Evans, J., Brannon, L., (1976). Understanding sentences with relative clauses. *Memory and Cognition* 4, 283–296.

Huang, C.-T. James. (1982). *Logical relations in Chinese and the theory of grammar*. Ph.D. dissertation, Massachusetts Institute of Technology.

Hulk, A. (1996). "The Syntax of Wh-Questions in Child French." *Amsterdam Series in Child Language Development* 5: 129-172.

Hulk, A. and Zuckerman, S. (2000). "The interaction between input and economy: Acquiring optionality in French Wh-questions." In *Proceedings of the 24th Annual Boston University Conference on Language Development [BUCLD 24]*, S. Catherine Howell, Sarah A. Fish, and Thea Keith-Lucas (eds), 438-449. Somerville, MA: Cascadilla Press.

Jakubowicz, C. (2004). Question formation in French SLI: Is movement costly? Paper

presented at *Groningen Workshop on the L1 Acquisition of Tense, Aspect, and Questions,* June 11, 2004, University of Groningen.

Jakubowicz, C. (2005) The Language Faculty: (Ab)normal Development and Interface Constraints. Paper presented at *GALA 2005*.

Jakubowicz, C. (2011). Measuring derivational complexity: New evidence from typically developing and SLI learners of L1-French. *Lingua 121*, 339-351.

Johnson, V.E., J. de Villiers, and H. Seymour (2005). Agreement without understanding? The case of third person singular –s. *First Language 25*(3), 317-330.

Labelle, M. (1990). "Predication, Wh-Movement, and the Development of Relative Clauses." *Language Acquisition* 1: 95-119.

Labelle, M. (1996). "The Acquisition of Relative Clauses: Movement or No Movement?" *Language Acquisition* 5: 65-82.

Lord & Paul (1997). Language and communication in autism. In D. J. Cohen & P. R. Volkmar (eds), *Handbook of autism and pervasive developmental disorders* (pp. 195–225). New York: Wiley.

Mathieu, E. (1999). WH in situ and the intervention effect. In *UCL Working Papers in Linguistics,* Corinne Iten and Ad Neeleman (eds), 441-472. London: University College.

McDaniel, D., McKee, C. & Bernstein, J. B. (1998). How children's relatives solve a problem for minimalism. *Language 74*, 308–34.

Miller, K. and C. Schmitt (2009). Variable vs. consistent input: Comprehension of plural morphology and verbal agreement in children. In J. M. Brucart, A. Gavarro and J Sola (Eds.), *Merging Features: Computation, Interpretation and Acquisition*. Oxford University Press.

36

Perez-Leroux, A. T. (2005). Number problems in children. In C. Gurski (Ed.),
    *Proceedings of the 2005 Canadian Linguistics Association Annual
    Conference*.

Perovic, A., N. Modyanova, E. Hanson, C. Nelson and K. Wexler. (2007).
    Investigations of language in autism: Evidence for a grammatical deficiency.
    Poster presented at Autism Research in the UK: From Diagnosis to
    Intervention. The Open University, UK.

Perovic, A., N. Modyanova., & K. Wexler. (2013). Comprehension of reflexive and
    personal pronouns in children with autism: A syntactic or pragmatic deficit?
    Applied Psycholinguistics 34, 813-835.

Plunkett, B. 1999. "Targeting complex structure in French questions." In *Proceedings
    of the 23rd Annual Boston University Conference on Language Development
    [BUCLD 23]*, Annabel, Greenhill, Heather Littlefield, and Cheryl Tano (eds),
    764-775. Somerville MA: Cascadilla Press.

Reinhart, T. (1997) Quantifier scope: How labor is divided between QR and choice
    functions. *Linguistics and Philosophy* 20:335-397.

Riches, N. G., Loucas, T., Baird, G., Charman, T. and 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*, 47-60.

Rizzi, L. (1990). *Relativized Minimality*. MIT Press, Cambridge, MA.

Rizzi, L. (2004). Locality and the Left Periphery. In *Structure and beyond*, A. Belletti
    (Ed.), Oxford University Press, 223–251.

Rizzi, L. (2013). Locality. *Lingua 130*. 169-186.

Roth, P. F. (1984). Accelerating language learning in young children. *Journal of

Child Language 11, 89–107.

Sheldon, A. (1974). The role of parallel function in the acquisition of relative clauses in English. *Journal of Verbal Learning and Verbal Behavior 13*, 272–281.

Shlonsky, Ur. (2012). Notes on wh in situ in French. In *Functional heads: The Cartography of Syntactic Structures 7*, Laura Bruge, Anna Cardinaletti, Giuliana Giusti, Nicola Munaro, and Cecilia Poletto (eds), 242-252. New York: Oxford University Press.

Starke, M. (2001). *Move Dissolves into Merge: a Theory of Locality*. Unpublished doctoral dissertation. University of Geneva.

Strik, N. (2008) *Syntaxe et acquisition des phrases interrogatives en français et en néerlandais :une étude contrastive.* Unpublished doctoral dissertation, Université Paris 8.

Tager-Flusberg, H., Calkins, S., Nolin, T., Baumberger, T., Anderson, M., & Chadwick-Dias, A. (1990). A longitudinal study of language acquisition in autistic and Down syndrome children. *Journal of Autism and Developmental Disorders 20*, 1–21.

Tager-Flusberg H., (2004) Strategies for conducting research on language in autism. Journal of Autism and Developmental Disorders. 2004;34:75–80.

Terzi, A, Marinis, T, Kotsopoulou, A, & Francis, K., (2014) Grammatical Abilities of Greek-Speaking Children with Autism, *Language Acquisition 21*, 4-44.

van Kampen, J. (1997). *First steps in wh-movement*. Delf: Eburon.

Villata, S., Rizzi, L., Franck, J. (in press). Intervention effects and Relativized Minimality: New experimental evidence from graded judgments. *Lingua.*

Zebib, R., Tuller, L., Prévost, P., Morin, E. (2013). *Formal Language Impairment in French-Speaking Children with ASD: A Comparative ASD/SLI Study*.

Advances in Language Acquisition. Newcastle: Cambridge Scholars
Publisher.

Zuckerman, S. (2001). *The acquisition of "optional" movement*. Unpublished
doctoral dissertation, University of Groningen.

Tables

Table 1. Levels of complexity based on Movement, Intervention, and Similarity

	Movement	Intervention	Similarity
Level 0	-	-	-
Level 1	+	-	-
Level 2	+	+	-
Level 3	+	+	+

Table 2. TD groups

		4-year old TD children [N=15]	6-year old TD children [N=15]	8-year old TD children [N=15]
Age	Mean	4.97	6.85	8.74
	Range	4.56-5.26	6.47-7.41	8.3-9.23
	SD	0.24	0.29	0.29

Table 3. Matched children with ASD and TD children

		Children with ASD [N=20]	TD children [N=19]
Age	Mean	9.53	7.59
	Range	5.5-16	4.89-9.23
	SD	3.22	1.5
Raven's matrices	Mean	24.93	26.21
	Range	9-36	14-34
	SD	7.72	6.35

Table 4. Structures tested across the four complexity levels

Level 0	Examples
Object questions in situ, –NP object	*...les éléphants poussent qui ?*
	... the elephants are pushing who?
Object questions in situ, +NP object	*...les éléphants poussent quel ours ?*
	... the elephants are pushing which bear?

Level 1	
Subject questions, -NP subject	*... qui__ pousse les éléphants ?*
	... who __ is pushing the elephants?
Subject questions, +NP subject	*... quel ours__ pousse les éléphants ?*
	... which bear __ is pushing the elephants?
Subject relatives, +NP subject	*... l'ours qui__ pousse les éléphants*
	... the bear who __ is pushing the elephants

Level 2	
Object questions ex situ, -NP subject	*... qui les éléphants poussent__?*
	... who the elephants are pushing __?

Level 3	
Object questions ex situ, +NP subject	*... quel ours les éléphants poussent__?*
	... which bear the elephants are pushing __.
Object relatives, +NP subject	*... l'ours que les éléphants poussent__*
	... the bear that the elephants are pushing __

Table 5. Mean accuracy (in percentage) and standard deviation by structure

	4-year olds	6-year olds	8-year olds
Level 0			
Object questions in situ, –NP object	87.8 (13.3)	97.8 (5.9)	100 (0)
Object questions in situ, +NP object	84.4 (14.7)	97.8 (5.9)	100 (0)
Level 1			
Subject questions, -NP subject	84.4 (11.7)	94.4 (7.5)	95.6 (12.1)
Subject questions, +NP subject	87.8 (10.4)	93.3 (7.2)	95 (7.6)
Subject relatives, +NP subject	88.3 (10.4)	97.2 (6.8)	97.2 (5.1)
Level 2			
Object questions ex situ, -NP subject	70 (20.1)	73.3 (19.7)	90 (13.8)
Level 3			
Object questions ex situ, +NP subject	60 (21.6)	78.9 (19.4)	88.9 (13.6)
Object relatives, +NP subject	45.6 (22.2)	82.8 (15.3)	86.7 (10.4)

Table 6. Mean accuracy (in percentage) and standard deviation by structure

	Children with ASD	TD children
Level 0		
Object questions in situ, –NP object	88.9 (21.5)	98.8 (4.5)
Object questions in situ, +NP object	88.9 (18.6)	96.4 (9.7)
Level 1		
Subject questions, -NP subject	77.2 (30.1)	95.3 (7.1)
Subject questions, +NP subject	77.2 (22.4)	95.2 (7.8)
Subject relatives, +NP subject	82.8 (20.3)	96.4 (7.1)
Level 2		
Object questions ex situ, -NP subject	66.7 (31.5)	79.8 (17.5)
Level 3		
Object questions ex situ, +NP subject	57.8 (37.2)	83.3 (21.7)
Object relatives, +NP subject	56.7 (31.8)	75.6 (20.3)

Table 7. Correlations between age/non-verbal abilities and performance on the experimental task

	Children with ASD		TD children	
	r	p	r	p
Age and Non-verbal IQ	0.253	0.362	0.866***	< 0.001
Age and Level 0	-0.112	0.692	0.650*	0.012
Age and Level 1	-0.074	0.792	0.583*	0.029
Age and Level 2	0.534*	0.04	0.266	0.357
Age and Level 3	0.37	0.175	0.738**	0.003
Non-verbal IQ and Level 0	0.512	0.051	0.614*	0.019
Non-verbal IQ and Level 1	0.621*	0.014	0.668**	0.009
Non-verbal IQ and Level 2	0.568*	0.027	0.468	0.091
Non-verbal IQ and Level 3	0.564*	0.028	0.568*	0.034

$* = p < 0.05$

$** = p < 0.01$

$*** = p < 0.001$

Figures

Figure 1. Example from the sentence-picture matching task
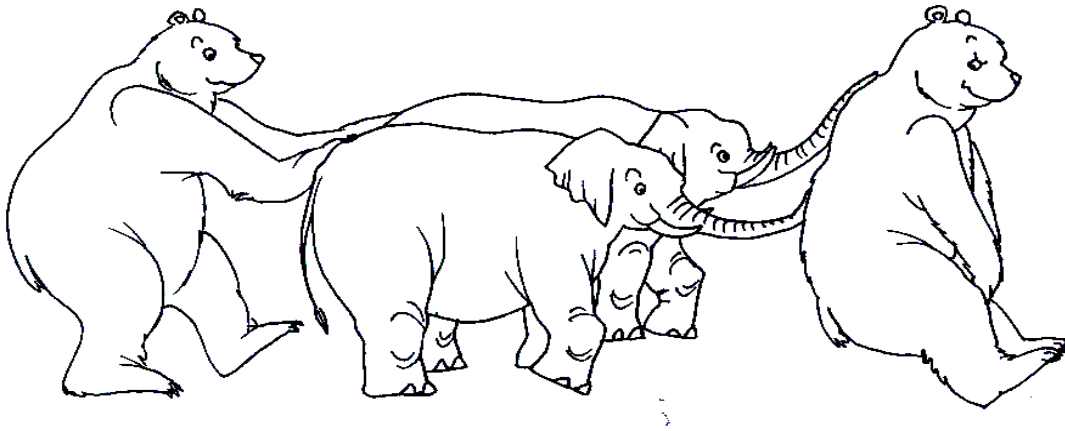
Figure 2: Accuracy in 4-, 6-, and 8-year old TD children

Error bars: 95% CI
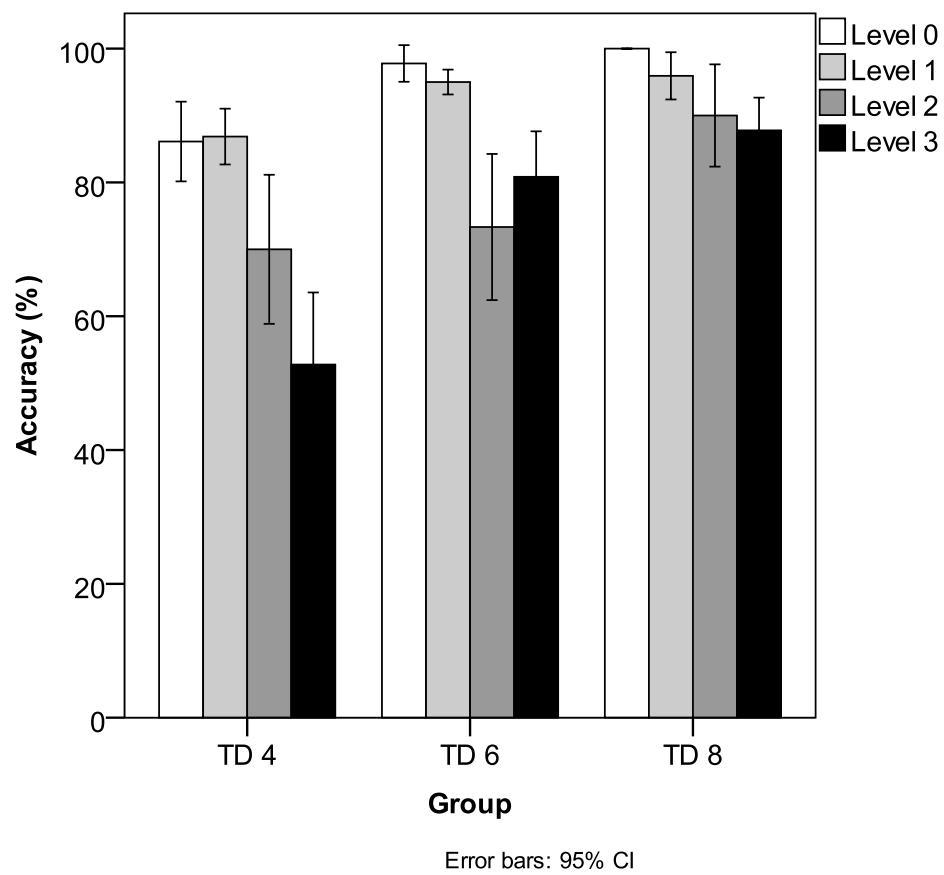
Figure 3: Structures with and without movement in TD children
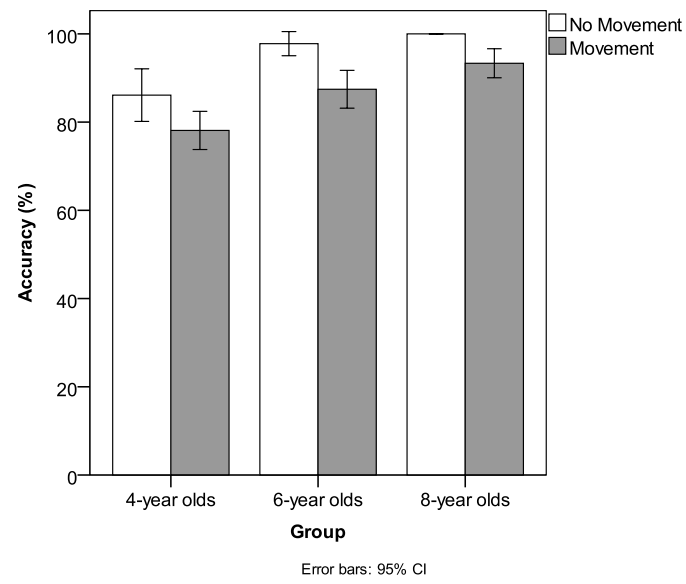

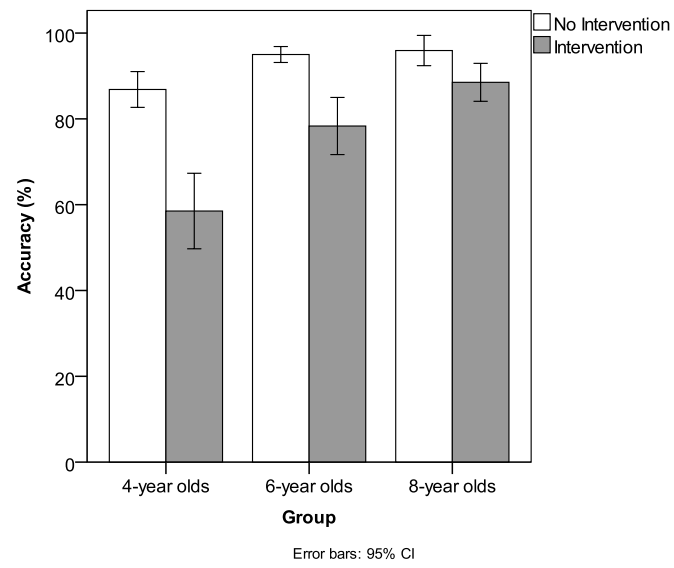Figure 4: Structures with and without intervention in TD children

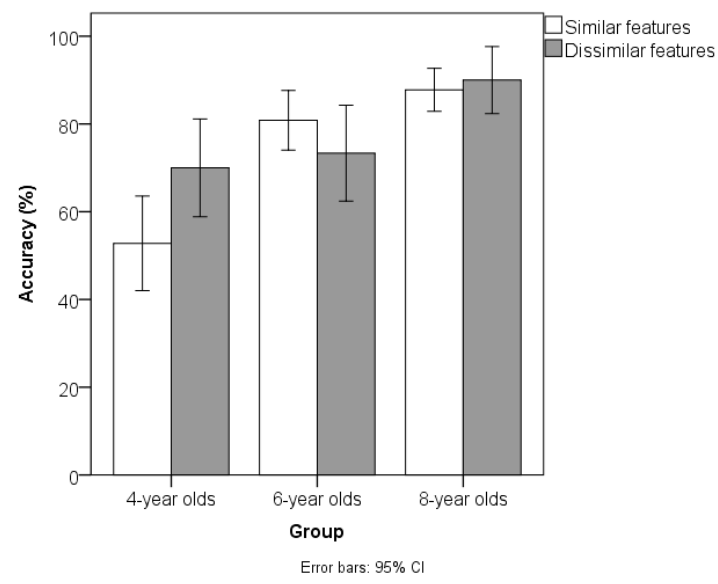Figure 5: Structures with and without feature similarity in TD children

47

Figure 6: Accuracy in children with ASD compared to their TD controls
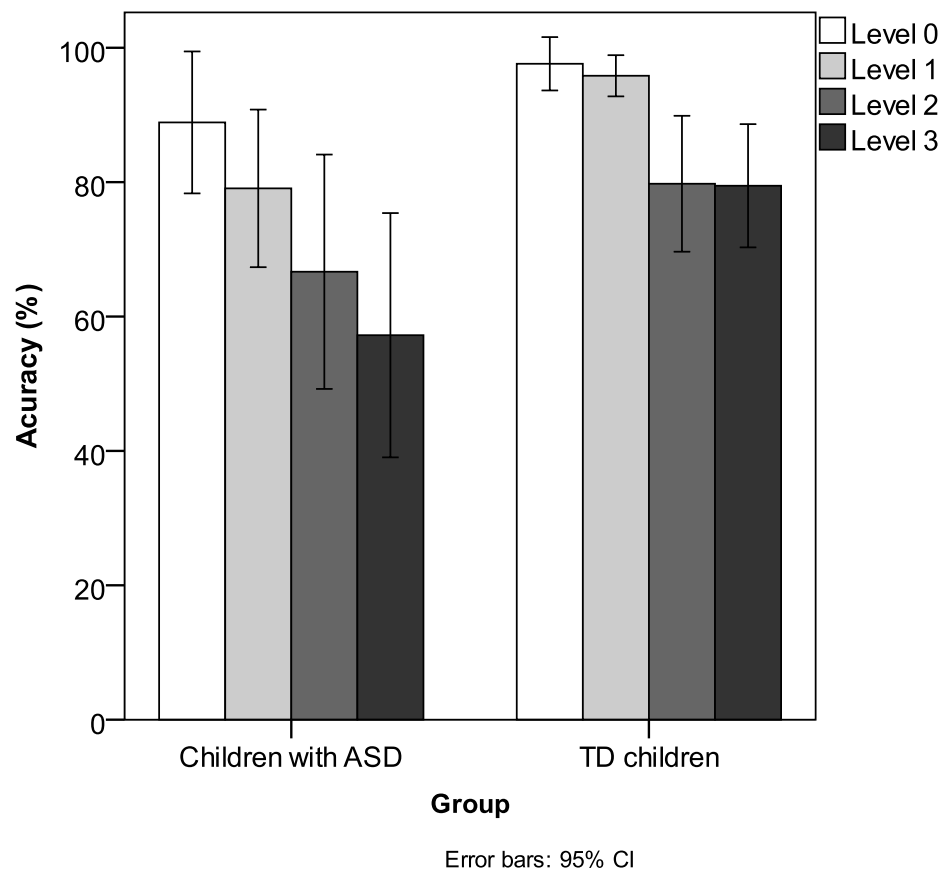
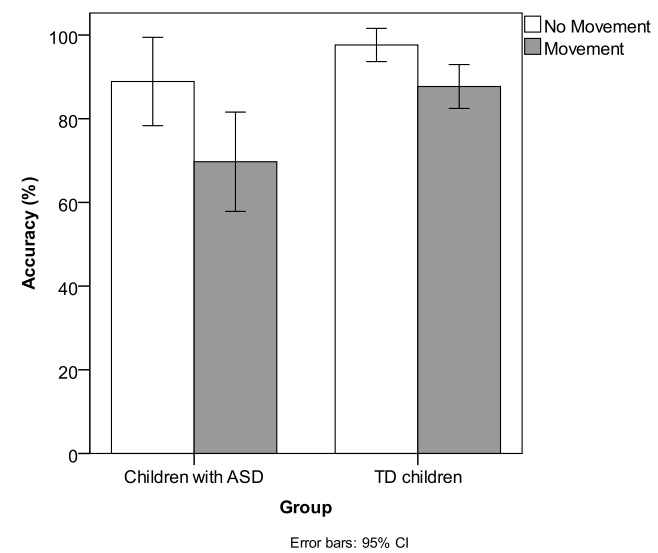Figure 7: Structures with and without movement in children with ASD vs. TD controls

Error bars: 95% CI

Figure 8: Structures with and without intervention in children with ASD vs. TD controls
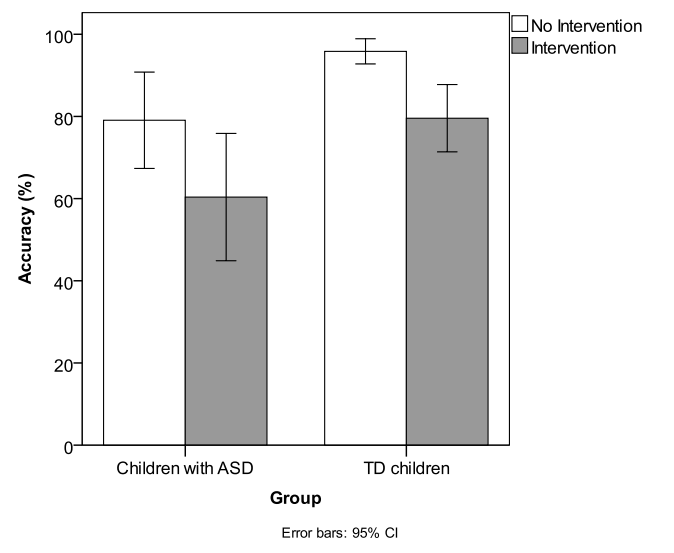

Error bars: 95% CI

Figure 9: Structures with and without feature similarity in children with ASD vs. TD controls

Error bars: 95% CI