UNIVERSITÉ DE TOURS

ÉCOLE DOCTORALE Santé, Sciences Biologiques et Chimie du Vivant

INSERM UMR 1253, Imagerie et Cerveau
Equipe 'Psychiatrie Neuro-fonctionnelle'

THÈSE présentée par :

Silvia SILLERESI

soutenue le: **10 décembre 2018**

pour obtenir le grade de : **Docteur de l'Université de Tours**

Discipline/ Spécialité: Psycholinguistique

Structural Language and Nonverbal Ability Profiles in Monolingual and Bilingual Children with ASD

THÈSE dirigée par :

Mme TULLER Laurie	Professeur, Université de Tours
M. PRÉVOST Philippe	Professeur, Université de Tours

RAPPORTEURS :

Mme GUASTI Maria Teresa	Professeur, Université Milano-Bicocca, Italie
Mme TSIMPLI Ianthi Maria	Professeur, Université Cambridge, GB

JURY :

M. BOUVARD Manuel	PU-PH, Centre Hospitalier Charles Perrens, Bordeaux et Université de Bordeaux
Mme GUASTI Maria Teresa	Professeur, Université de Milano-Bicocca, Italie
Mme LATINUS Marianne	Chercheur associé INSERM, UMR 1253, Tours
M. PRÉVOST Philippe	Professeur, Université de Tours
Mme TSIMPLI Ianthi Maria	Professeur, Université de Cambridge, GB
Mme TULLER Laurie	Professeur, Université de Tours

To my parents Sandra and Roberto,
who taught me that when a beautiful thing ends
is because there is another waiting for you.

Acknowledgements

First of all I would like to express my deepest gratitude to my supervisors Laurie Tuller and Philippe Prévost, for their continuous support, for their patience, motivation, and immense knowledge. In these three years I had the chance to learn from you how to make the first steps to becoming a researcher and I will never thank you enough for this. Your kindness and your incredible ability in teaching will always be some of the greatest memories of this experience. But most importantly, I would always be grateful to you for sharing your deep love for this field of research. This was the biggest gift of all.
Someone once told me that you can consider yourself lucky if, in your life, you come across a mentor. Well, I was lucky enough to find two. I hope this is just the beginning.

I would like to thank all the members of my thesis committee, Maria Teresa Guasti, Ianthi Maria Tsimpli, Manuel Bouvard and Marianne Latinus, for taking the time to read my work and for their insightful comments.

To all the amazing people in the linguistic department, Racha Zebib, Sandrine Ferré, Christophe dos Santos, Nicola Lampitelli, Eléonore Morin, Laetitia de Almeida. Thank you for all the precious advices, for always willing to help and lend kind words through this whole journey. It would not have been the same without you.

I am also deeply thankful to Pr. Bonnet-Brilhault who allowed me to be part of the wonderful family of the Equipe "Psychiatrie Neuro-fonctionelle" at the hospital, to all the doctors who were involved in the research protocol, Agathe Saby, Marie Guimard-Brunault, Pasquale Glorioso and especially to Joëlle Malvy who coordinated with patience all the recruitments through these years.
Thanks to all the children and the families who agreed to participate in the study.

A big thank goes to all the speech language pathologists, Magalie Bataille, Marine Roquebert, Sarah Boucher Bonheur for their help and their suggestions and especially to Chloé Chasseloup who helped me with the assessments in my first year and to Nathalie Courtois, whose experience and knowledge helped me in several occasions.

I would also like to thank all the educators at the day-care hospital, the psychologists and who helped me in collecting data, in particular Patrice Gillet, Romuald Blanc, Julie Allard-

Ech-Chouikh, Cindy Le Menn Tripi and all the staff of CHRU Bretonneau. A very special thank you goes to Jacqueline Sapede who always welcomed me with a smile.

To all the wonderful people of the "premiere étage en pedospychiatrie": Sylvie Roux, Remy Magné, Benedicte Schubart, Céline Courtin, Nathalie Bourron, Luce Corneau, Clotilde Gronnier, Marie Gomot, Alix Thillay. Thank you for always being so nice to me and for making the office a very nice place to work.

To my partners "in crime" Céline David, Judith Charpentier, Klara Kovarski and Elsa Legris, who shared with me the beauty and the difficulties of all these years of hard work. Thank you for your help and constant presence. You made this journey much more fun and definitely less stressful. It would not have been the same without you *les filles.* Thanks also to Elodie Léger, Nufar Sukenik, Thypanie Prince and Bénédicte Grandon for the precious advices and to Mathilde Chailleux and Samantha Ruvoletto for your friendship through these years.

I owe a big thank you to all the researchers I had the lucky to encounter in this journey and who were kind enough to share with me some of their knowledge, Janette Schaeffer, Chiara Cantiani and Cinzia Pieraccini.

When I came to Tours three years ago I would not expect to find another family. Thank you Donatello and Diane for making me feel at home and for let me be part of your beautiful family. A big hug goes to all my friends in Poitiers, Niort and Paris who helped me every step of the way.

I must change to Italian for this last part.

Un grande grazie va a tutti i miei amici in Italia che mi hanno accompagnato anche a kilometri di distanza. Grazie per avere reso questa fatica un po' più leggera.

Ai miei nonni che hanno sempre creduto in me.

Ai miei genitori, grazie per esserci stati in questo ennesimo lungo viaggio e per aver supportato tutte le mie decisioni. Niente di tutto questo sarebbe stato possibile senza di voi. Grazie per avermi educato ad essere la persona che sono oggi.

A Pietro, grazie per l'amore, la pazienza e la tenerezza con cui mi stai vicino ogni giorno. Ma soprattutto grazie per aver abbracciato il mio Destino ed averlo fatto diventare il nostro.

Résumé

Le diagnostic de Trouble du Spectre Autistique (TSA) prévoit une spécification pour toutes les occurrences de trouble du langage et/ou déficience intellectuelle. En ce qui concerne les troubles du langage, plusieurs études ont montré que chez les enfants avec TSA il existe un sous-groupe d'enfants qui manifestent des troubles sévères du langage structurel, en morphosyntaxe et phonologie (TSA-TL) similaires à ceux rencontrés chez les enfants avec Trouble Spécifique du Langage (TSL), alors que le reste des enfants présentent des habilités langagières dans la norme (TSA-LN), comme les enfants à développement typique (DT). Concernant les déficits cognitifs, les études actuelles qui ont une large popualtion ont mis en évidence que près d'un tiers des individus appartenant au spectre présentent une déficience intellectuelle.

Néanmoins, très peu d'études ont exploré de façon explicite les combinaisons possibles entre (dys)fonctionnements langagiers et (dys)fonctionnements cognitifs avec pour but de mieux définir les profils de langage structurel et de développement cognitif dans le TSA. Notre étude a proposé une investigation systématique de chacune de ces habilités basée sur des mesures clairement motivées, chez un groupe de 51 enfants monolingues (n = 37) et bilingues (n = 14) avec TSA, âgés de 6 à 12 ans. Nous avons argumenté en faveur de l'utilisation de deux tâches de répétition créées spécifiquement pour mesurer la complexité linguistique, en morphosyntaxe, tâche de répétition de phrases (SR) et en phonologie, tâche de répétition de non-mots (NWR) et de trois taches non-verbales, les Matrices Progressives de Raven (RPM) et deux subtests de l'Echelle Wechsler, Cubes et Matrices pour évaluer les capacités cognitives des enfants avec TSA. Nous avons utilisé une méthode de classification non supervisée, l'analyse en cluster, pour identifier les profils de langage structurel et les habilités non-verbales chez les enfants monolingues avec TSA. Successivement, nous avons comparé les réalisations phénotypiques langagières de profils obtenus avec deux groupes contrôles, pour déterminer si les capacités des enfants avec un trouble du langage ressemblaient à celles des enfants avec TSL (*n* = 26) du même âge, et si les enfants avec des habilités langagières dans la norme se comportaient comme les enfants DT âgés de 4 à 12 ans, comme précédemment décrit dans la littérature. Enfin, nous avons intégré à la précédente analyse les enfants bilingues afin de vérifier si l'exposition à une deuxième langue pouvait causer l'émergence de profils d'habilités différents de ceux identifiés chez les enfants monolingues avec TSA.

Trois constats importants se dégagent de notre étude. L'analyse en cluster a suggéré

l'existence de cinq profils possibles chez les enfants monolingues et bilingues avec TSA : trois profils ayant des habilités langagières dans la norme et deux profils ayant des troubles langagiers sévères. Parmi ces profils, nous avons mis en évidence les quatre profils logiquement possibles à partir de la combinaison des habilités structurelles du langage et des habilités cognitives. De façon importante, l'existence de profils discordants comme un TSA-TL avec un QI non-verbal dans la norme et un TSA-LN avec un QI non-verbal bas, suggère que les enfants monolingues comme les enfants bilingues peuvent en effet manifester des troubles du langage en présence d'une intelligence non-verbale normale ou des habilités langagières dans la norme en présence d'une intelligence non-verbale déficitaire, ce qui renforce l'hypothèse de l'existence d'un module séparé pour le langage dans le cerveau.

L'analyse de réalisations phénotypiques langagières de cinq profils via la performance globale et une analyse approfondie et qualitative des erreurs, a démontré que les deux groupes d'enfants TSA-TL ont montré un déficit mophosyntaxique beaucoup plus sévère que les enfants avec TSL, alors que leur performance ne différait pas sur le plan phonologique. De même, les trois groupes d'enfants avec un profil TSA-LN ressemblaient aux enfants à DT uniquement en ce qui concerne leur performance en phonologie, tandis que certains enfants TSA-LN (deux des trois profils) montraient quelques difficultés sur les structures syntaxiquement plus complexes. Ces résultats questionnent la légitimité des hypothèses qui supposent l'existence d'un profil TSL chez les enfants avec TSA et la nature « normale » des habilités structurelles du langage chez les enfants avec TSA-LN.

Enfin, l'analyse des enfants bilingues avec TSA a démontré que les mêmes profils identifiés chez les enfants monolingues étaient présents chez les enfants exposés à deux langues, ce qui indique que le bilinguisme ce n'est pas un facteur préjudiciable pour les enfants avec TSA. Ce qui a déjà été démontré chez les enfants avec TSL et chez des enfants avec le Syndrome de Down également.

En conclusion, notre étude atteste qu'une compréhension des profils de langage structurel et des habilités non-verbales chez les enfants monolingues et bilingues avec autisme peut être permise uniquement via l'utilisation de mesures appropriées (qui permettent une analyse quantitative et qualitative des erreurs) et via une exploration à travers le spectre. Du point de vue clinique, une division en sous-types pourrait augmenter la possibilité de créer des traitements adaptés aux besoins spécifiques des individus, en se basant sur leurs points forts et leurs points faibles.

Abstract

A diagnosis of Autism Spectrum Disorder (ASD) includes specification of any co-occurrence with language impairment and/or cognitive disabilities. Regarding language impairment, studies have reported that among verbal children with ASD a subgroup of children manifests significant structural language impairment (ASD-LI), similar to the one displayed by children with Specific Language Impairment (SLI), while the rest display normal abilities (ASD-LN), like those found in Typically Developing (TD) children. Regarding cognitive disabilities, current large-scale studies have found that roughly one-third of individuals on the autism spectrum are affected.

However, few studies have explicitly explored the possible combinations of language (dis)ability and cognitive (dis)ability, with the aim of better defining profiles of structural language and cognitive development in ASD. Our study proposed a systematic investigation of these abilities in fifty-one 6- to 12-year-old monolingual (n = 37) and bilingual (n = 14) children with ASD based on explicitly motivated measures. We argued for the use of two linguistically based repetition tasks, Sentence Repetition (SR) and Nonword Repetition (NWR), for evaluating morphosyntax and phonology, and three nonverbal (NV) tasks, Raven's Progressive Matrices (RPM), Block Design and Matrix Reasoning (two subtests of the Wechsler Scale) for evaluating cognitive abilities in children with ASD. An unsupervised machine learning approach, cluster analysis, was used to identify the profiles of structural language and NV abilities in monolingual children with ASD. Then, a comparison of the phenotypical linguistic realisation of the profiles that emerged from this analysis was run with two control populations in order to determine whether the children with ASD presenting language impairment would resemble age-matched children with Specific Language Impairment (SLI) (n = 26) and whether the children with ASD displaying normal language abilities would resemble TD children (4- to 12- year-old, n = 84), as suggested by previous studies in the literature. Finally, the bilingual children with ASD were integrated into the previous analyses with the intent of verifying whether being exposed to a second language would cause the emergence of different profiles of structural language / NV abilities compared to monolingual peers.

Three main results emerged from our study. The cluster analysis suggested the existence of five possible profiles in both monolingual and bilingual children with ASD: three profiles with normal language abilities and two profiles displaying language impairment. Among these five profiles, all four logically possible structural language / NV

abilities combinations were detected. Crucially, the existence of discrepant profiles of abilities, the ASD-LI with average NVIQ profile and the ASD-LN with low NVIQ profile suggests that both monolingual and bilingual children with ASD can display impaired language abilities in presence of spared nonverbal intelligence or spared language abilities in the presence of impaired nonverbal intelligence, reinforcing the hypothesis of the existence of a separate language module in the brain.

The analysis of the phenotypical realisation of language abilities in the five profiles via overall performance and in depth qualitative error analysis showed that both groups of children with ASD-LI showed much more severe impairment in morphosyntax than the children with SLI, while their impairment did not differ for phonological abilities. Similarly, the three groups of children with ASD-LN showed similar phenotypical profiles of phonological abilities as age-matched TD children, while for two of the three LN profiles a selective drop in performance was observed on complex syntactic structures. These results question the legitimacy of the hypothesis that there is an SLI profile in ASD and that children with ASD-LN display spared structural language abilities.

Finally, the analysis of bilingual children with ASD showed that the same profiles of abilities as the ones found in monolingual children with ASD could be identified in children exposed to more than one language, indicating that bilingualism is not a detrimental factor for children with ASD, similarly to what has been demonstrated for other clinical populations, notably SLI and Down Syndrome.

Our study provides evidence that understanding structural language/nonverbal ability profiles in both monolingual and bilingual children can only be achieved through the use of proper tools (that enable both quantitative and qualitative in depth error analyses) and through wide investigation across the entire spectrum. Clinically speaking, subtyping would increase the possibility of developing treatments tailored to the specific needs of individuals, based on their particular pattern of strengths and impairment.

Table of Contents

List of Tables

List of Figures

List of Appendix

List of Most Frequent Acronyms and Abbreviations

ADI-R	Autism Diagnostic Interview-Revised
ADOS	Autism Diagnostic Observation Schedule
APA	American Psychiatric Association
AS	Asperger's Syndrome
ASD	Autism Spectrum Disorder
ASD-LI	Autism Spectrum Disorder with Language Impairment
ASD-LN	Autism Spectrum Disorder with Normal Language
BILO	Bilans Informatisés de Langage Oral
CARS	Childhood Autism Rating Scale
CDC	Centre for Disease Control and Prevention
CELF	Clinical Evaluation of Language Fundamentals
DSM	Diagnostic and Statistical Manual of Mental Disorders
ECA-R	Evaluation of Autistic Behaviours
EDEI-R	Echelles Differentielles d'Efficience Intellectuelle
ELO	Evaluation de Langage Oral
EVIP	Échelle de Vocabulaire en Images Peabody
FSIQ	Full Scale Intelligence Quotient
HF	High Functioning
HFA	High Functioning Autism
ICD	International Classification of Diseases
LF	Low Functioning
LFA	High Functioning Autism
LITMUS	Language Impairment Testing in a Multilingual Setting
MorsynP	Expressive Morphosyntax
MorsynR	Receptive Morphosyntax
MSEL	Mullen Scales of Early Learning
NVIQ	Non Verbal Intelligence Quotient
NWR	Nonword Repetition

PDD-NOS	Pervasive Developmental Disorder Not Otherwise Specified
PPVT	Peabody Picture Vocabulary Test
PRI	Perceptual Reasoning Index
PSI	Processing Speed Index
RPM	Raven Progressive Matrices
SLI	Specific Language Impairment
SLP	Speech-Language Pathologists
SR	Sentence Repetition
TD	Typically Developing
VCI	Verbal Comprehension Index
VocP	Expressive Vocabulary
VocR	Receptive Vocabulary
WISC	Wechsler Intelligence Scale for Children
WMI	Working Memory Index

Part I

Introduction

Overview

Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder, characterized and diagnosed by impairments in social communication and social interaction in the presence of restricted, repetitive behaviours or interests (DSM-5 American Psychiatric Association, APA, 2013). It has a strong genetic basis, although the genetics of autism are complex and it is unclear whether ASD is explained more by rare mutations or by rare combinations of common genetic variants (Talkowski et al., 2014). Current population prevalence is estimated at ~1.5% in developed countries around the world (Baio et al., 2018; Baxter et al., 2015; Lyall et al., 2017), and is about 4.5 times more common among boys than among girls. Parental concerns usually emerge in the first three years of life of the child because of self-evident delays in social interaction (abnormalities in eye contact and body language or deficits in understanding and use of gestures; lack of facial expressions and nonverbal communication, failure in showing or pointing to objects of interest, a.o.) and in language onset and development (failure to call their parents by name, lack of coordinating vocalizations with their intentions, failure of normal back-and-forth conversation, a.o.).

Autism is a highly heterogeneous disorder: no two individuals with ASD show exactly the same profile and its description becomes even more complicated when we look at the disorder from a chronological and developmental perspective. The term *spectrum* was introduced by the DSM-5 as an "umbrella term" in order to include the wide range of symptoms, capacities, and levels of impairments that individuals with ASD may display. Symptoms can range from mild to severe and often change over time, which makes it even more important to rely on a core and consistent system of diagnosis (McPartland et al., 2012). For these reasons, in 2013, the APA suggested replacement of the autism triad of the DSM-IV, given in (1), with a simplified autism dyad in the new DSM-5, comprising the criteria in (2), commonly referred to as the "first dimension" and the "second dimension".

(1) a. Qualitative impairments in social interaction;
b. Qualitative impairments in communication;
c. Restricted, repetitive and stereotyped patterns of behaviour, interests and activities

(2) a. Persistent deficits in social communication and social interaction across multiple contexts, (combination of the first two previous criteria);
b. Restricted, repetitive patterns of behaviour, interests, or activities, (RRB).

To be diagnosed with ASD, children must show deficits in all three domains of the social communication impairment dimension (2a): 1) social-emotional reciprocity, 2) nonverbal communicative behaviour, and 3) developing, maintaining, and understanding relationships; and on at least two of the four criteria related to the restrictive and repetitive behaviours dimension (2b): 1) stereotyped or repetitive motor movements, use of objects, or speech, 2) insistence on sameness, inflexible adherence to routines, or ritualized patterns or verbal nonverbal behaviour, 3) highly restricted, fixated interests that are abnormal in intensity or focus, and 4) hyper- or hyporeactivity to sensory input or unusual interests in sensory aspects of the environment (APA, 2013). With the introduction of the DSM-5, a diagnosis of ASD now subsumes the previous labels (DSM-IV) of Autistic Disorder (i.e. classical Kanner's Autism), Asperger's Syndrome (AS) and Pervasive Developmental Disorder Not Otherwise Specified (PDD-NOS). In May 2018, the International Classification of Diseases, ICD-11 (WHO, 2018) mirrored the criteria of the DSM-5 introducing into its classificatory system the idea of a spectrum of disorders for autism.

The change to a "spectrum" was not only motivated by the heterogeneity of autistic symptomatology. It was also motivated by the fact that the variety of profiles in children with autism is diverse since ASD is almost always accompanied by one or more co-occurring conditions that may develop during different phases of the life span. These conditions include language impairment, intellectual disability, medical or genetic conditions (epilepsy, sleep or gastrointestinal problems), other neurodevelopmental, mental or behavioural conditions (attention-deficit/hyperactivity disorder, tic disorders, anxiety disorder, coordination disorder, disruptive behaviour disorder, conduct disorder, bipolar disorder, Tourette's disorder, self-injury, feeding disorder, elimination disorder, sleep disorder and/or depression), and catatonia (APA, 2013). For the present study we will focus on language impairment and intellectual disability.[1]

When a child is diagnosed with ASD, the DSM-5 requires specification of any accompanying language and/or intellectual impairment. But what does each of these refer to? There is little current consensus about the definition of either of these properties.

Concerning language impairment, besides the universally recognized deficit in pragmatics identified as a diagnostic criterion in the first dimension of the DSM-5 (2a),

---

[1] Children presenting other co-occurring conditions, especially medical or genetic conditions and other neurodevelopmental disorders, were excluded from the present study. This choice was made on the consideration that these comordibities may cause too much variability in our population and might prevent reliable conclusions.

language abilities of children with ASD have typically been reported on the basis of vocabulary testing or developmental scales. However, this approach provides a very narrow picture of language capacities, entirely skipping over underlying structural language abilities (morphosyntax and phonology). Studies on verbal children with autism, approximately 70% of children on the spectrum (Tager-Flusberg & Kasari, 2013), have reported that 67%-75% of these children manifest mixed expressive/receptive language impairment (a group sometimes referred to as "autism with language impairment", ASD-LI), while the rest present normal language abilities (ASD-LN, autism with normal language) (Allen & Rapin, 1980; Tager-Flusbger, 2006).

Regarding intellectual impairment current large-scale studies have reported a mean rate of 31% of children with ASD, grounding their results on the general classification of intellectual disability at < 70 standard score on Full Scale IQ (FSIQ) (Center for Disease Control and Prevention (CDC), 2014). This has allowed researchers to separate children into profiles with High Functioning (HF) IQ ≥ 70 and Low Functioning (LF) IQ < 70. However, there is evidence that the evaluation of cognitive abilities through FSIQ scores is hard to interpret and it cannot be so neatly divide into profile of HF and LF capabilities because of the great heterogeneity displayed by children with ASD from one intellectual domain to the other. It is a well-known fact that children with ASD may display peaks of abilities in some areas of cognitive functioning (e.g. in the nonverbal, visuospatial domain) and valleys of performance in others (e.g. verbal abilities, working memory, processing speed) (see Nader et al., 2016 for an overview).

In the past decades, interest among researchers and clinicians in delineating different subgroups within the ASD population, based on independent consideration of each of these two comorbidities, has grown. So far, researchers have demonstrated that, along with different phenotypical realizations of autistic symptoms, linguistic and intellectual deficits may display multiple degrees of severity and impairment. However, very little is currently known about the interaction between these two domains.

The aim of the present study was to put the heterogeneity of ASD to the forefront by investigating whether clear profiles related to structural language and cognitive abilities emerged when investigation was extended to the entire spectrum. In order to explore this question, we explicitly argued for the use of specific measures of formal language abilities and of NV cognitive abilities in a group of 37 monolingual children with ASD aged 6- to 12-year-old. After considering possible links between these measures and factors previously argued to predict them (autism severity, age of first word, age of first sentence), we explored structural language / nonverbal ability profiles by using an unsupervised machine learning approach, cluster analysis, that took in consideration both linguistic and cognitive abilities of children with ASD. A comparison of the phenotypical linguistic realisation of profiles emerged from this analysis was run with control populations, in order to determine whether children presenting language impairment would resemble to age-matched children with Specific Language Impairment (SLI) and whether children displaying normal language abilities would resemble to Typically Developing (TD) children (4- to 12- year-old). Finally, moving from the results of our analysis of monolingual children with ASD, we investigated whether a group of 14 children with autism (aged 6- to 12- year-old) exposed to more than one language would display similar profiles of structural language / NV abilities as their monolingual peers. Although bilingualism is recognised as an increasing phenomenon worldwide and it has become an important topic of research for linguistic and cognitive abilities in children with developmental disorders (Kay-Raining Bird et al., 2016; Paradis, 2016) very few studies have investigated language and cognitive profiles of abilities in bilingual children with ASD. We aimed to start filling this shortcoming in the literature.

To our knowledge, this was the first study that investigated the interaction between structural language abilities (both phonology and morphosyntax) and cognitive abilities in both monolingual and bilingual children with ASD, describing the phenotypical characteristics of the emerged profiles.

The present study will be organised as follows: the first part (Part I) will provide background information, including a review of previous studies on both language and cognitive capacities of monolingual children with ASD, as well as a presentation of the theoretical framework of this study. In the first chapter we will report on current knowledge on language abilities in children with ASD. We will divide this literature review into two parts. The first part will report briefly on pragmatic impairment and the theories that have tried to account for it. The second part will focus on structural components of language, phonology and morphosyntax. We will describe previous research on structural language abilities in ASD and we will argue why formal language impairment should be considered as a primary domain of research in ASD. This review will lead us to hypothesize that two particular types of tasks, repetition of nonwords and sentences, should be considered the best tools for evaluating structural language abilities in children with ASD.

In the second chapter we will report on intellectual impairment and its relation with formal language abilities. We will focus on the evaluation of cognitive abilities in the autistic population and argue for the choice of nonverbal IQ (NVIQ) measures as the best suited measures for the evaluation of general fluid reasoning in ASD and the most reliable measures to be put in relation to linguistic abilities.

Finally in the third chapter we will investigate other factors that may be related to formal language or cognitive impairment in ASD. Notably we will look at whether generalized measures of severity of autism symptoms might be related with the performance of children on formal aspects of language and/or to intellectual disability and whether developmental factors such as age of first word and age of first sentence may have an impact in predicting outcomes of language and cognitive abilities in ASD.

At the end of the literature review, presented in the first three chapters, we will move to the experimental part of our work (Part II), after presenting the main research questions.

Chapter four will describe the method used to address the research questions of whether clear profiles of structural language / NV abilities would emerge from the use of specific measures for both these abilities.

In Chapter five we will explicitly argue for the use of specific measures of formal language abilities (two linguistically based repetition task for morphosyntax and phonology) and of NV cognitive abilities (tasks of fluid reasoning and visuospatial abilities) for evaluating children on the spectrum, via a comparison with standardized tests.

Then, in Chapter six we will consider possible links between these measures and factors previously argued to predict language and cognitive abilities (autism severity and developmental factors). Finally we will explore structural language / nonverbal ability profiles by using an unsupervised machine learning approach, the cluster analysis, which has the function to group data of similar kinds into respective categories.

In Chapter seven we will use the profiles of structural language/NV abilities emerged from Chapter six, to compare formal language abilities of monolingual children with ASD with those of monolingual children with SLI and both young and age-matched TD children, exploring in a quantitative, qualitative and developmental perspective the phenotypical realisation of performance of children with ASD on the two linguistically based repetition tasks, including in depth error analysis to determine intact versus impaired performance.

Part III will explore the applicability of the analysis proposed for monolingual children to a group of bilingual children with ASD. Chapter eight will briefly report on the very few studies that investigated language and cognitive abilities in bilingual children with ASD, highlighting the lack of studies on the topic, and it will present the applied methodology. Study results on bilingual children will be presented in chapter nine.

Finally Part IV will discuss general results and conclusions of both monolingual and bilingual studies in relation to previous literature.

We think that our study and the identification of clear profiles of structural language and NV abilities would help researchers creating more homogenous groupings that could facilitate pinpointing the different phenotypes of autism (e.g., linguistic capabilities and intellectual abilities) and for clinicians, subtyping would increase the possibility of developing treatments tailored to the specific needs of individuals, based on their particular pattern of strengths and impairment.

Chapter I
Structural language and Autism Spectrum Disorder

1.1 Introduction

The main purpose of this chapter was to propose a fine-grained analysis of current knowledge on linguistic abilities in children with ASD. The following literature review aimed at organizing the tangled *status quaestionis* about language abilities in autism (with a main focus on formal language). We will separate studies between those investigating structural language abilities via standardized tests and those that applied experimental tasks specifically constructed to evaluate specific constructions. At the end of the literature review we will conclude that there may be a better way to evaluate structural language abilities in this population, which is represented by the use of repetition tasks, sentence repetition for morphosyntax and nonword repetition for phonology, specifically constructed to evaluate language abilities.

1.2 Language characteristics in ASD

In the DSM–5, language impairment is no longer included as a core symptom of ASD, though clinicians are required to note whether or not a child presents a comorbid language disorder (APA, 2013). Communication deficits, such as social aspects of language use, have been universally reported and widely studied in children with ASD. This is the key feature that distinguishes communication impairment in autism from that of the great majority of children with Specific Language Impairment (SLI). SLI is a neurodevelopmental disorder that involves specific difficulties in mastering structural aspects of language, independently from any kind of intellectual disability, sensory impairment or neurological dysfunction. Diagnostic standards for SLI involve language achievement levels below cut-off values of age expectations and a NVIQ within the norms (Tomblin, 2011). Children must also be free from other developmental or sensory impairments, which include pervasive neurodevelopmental disorders, such as ASD. The aetiology of SLI currently remains unknown; however, in the last two decades there has been a growing body of evidence supporting a genetic cause.[2] The SLI label has been used since the 1980s, but has been recently questioned due to controversy about important aspects of the diagnosis (Reilly et al., 2014). The 'specific' element seemed not to reflect clinical realities (children with SLI do not all present with the same profile of difficulties) and excluded many children from clinical services. The CATALISE consortium (a multinational and multidisciplinary Delphi consensus study of problems of language development which dealt with the lack of agreement about criteria and terminology for children's language difficulties in research and clinical practice) suggested that the label "DLD", Developmental Language Disorder, should be used instead of SLI to avoid outdated assumptions about the causes of the child's difficulties (Bishop et al., 2017). Nevertheless, even if we are aware of the new guidelines, the present study will refer to this impairment as SLI, in conformity with the original label found in almost the totality of the studies reported in this literature review (which were run before the CATALISE action). This usage furthermore will guarantee continuity with the acronym "LI" in the ASD-LI label.

Starting with the earliest studies in autism, researchers have indicated that individuals with ASD systematically display serious pragmatic deficits (Baltaxe, 1977; Kanner, 1946; Tager-Flusberg et al., 1990). However less attention has been paid to formal aspects of language, which include domains such as phonology and morphosyntax. This can be explained in part by the fact that formal language impairment, although an important component of ASD, is not present in every individual with this condition, and is not usually considered necessary for a diagnosis (Tek et al., 2014). Nonetheless, delineating the nature of structural impairment is essential from the very first "linguistic steps" of these children, since delay in onset and/or subsequent development of language are one of the first concerns of parents of toddlers with autism. Moreover, the consequent therapeutic plan can be more tightly geared to the particular phenotypical impairment showed by the child: in other words each child can be treated on the basis of his/her specific disorder, whether it be only pragmatic or phonological / morphosyntactic in addition.

The general classification that is given in the first dimension of the DSM-5 under the umbrella term "communication deficits" (2a) is not even remotely sufficient to encompass the diversity of linguistic profiles in ASD. In fact, the nature of linguistic abilities in children with ASD shows great variation from one individual to another, ranging from the absence of

[2] One of the greatest revolutions for both SLI and ASD in the past decades, was the detachment from environmental explanations as their genetic / heritable character has become clear, particularly from twin studies (Bishop, 2001, 2006; Lindgren et al., 2009; Ronald & Hoekstra, 2011).

verbal ability or little functional communication, to well-developed syntactic capabilities and functional speech (Anderson et al., 2007; Tek et al., 2014). General, current reports indicate that among children with ASD about 30% are minimally verbal or entirely nonverbal (Anderson et al., 2007; Tager-Flusberg & Kasari, 2013). Verbal children should, therefore, represent the remaining 70% of the spectrum. However, the nature of linguistic deficits even in verbal children with ASD remains unclear: because of the large variation in language outcomes, there is not much consensus among researchers as to which aspects of language are intact or impaired (and why). For example, if the vast majority of individuals with ASD shows deficits in at least one formal language domain, some individuals, such as those with Asperger's syndrome (DSM-IV), currently reported to constitute 2.5% of the ASD population (CDC, 2014), do not show any shortcomings in phonology or morphosyntax.

How do we reconcile the heterogeneity of linguistic levels observed across the spectrum with diagnostic criteria and, more importantly, with intervention in clinical practice? The answer to this question relies in part in abandoning the practice of characterizing language through a single measure (such as "communicative deficits"), especially in autism, which may show selective impairments in one or more domain(s) that constitute language. A better approach would be to try and identify language profiles in these children through independent assessment of multiple language skills and compare them with patterns of performance in control groups. Constructing such individual linguistic profiles may contribute to the detection of patterns of strengths and weaknesses in multiple domains, which could lead to the identification of syndrome-specific features of language development in autism. In order to explore the different components of language and their general outcomes in ASD, the following sections will describe our current knowledge as reported in the literature on language and autism.

1.2.1 Pragmatics in ASD

The concept of "pragmatics" reflects the use of language as a tool for communication; it refers specifically to the ability to use language in contexts of social interaction. Eigsti et al. (2011: 683) pointed out that pragmatics is a very broad term used to refer to "[…] both linguistic functions, such as register […], negotiation of turn-taking, and the choice of referential expressions […], as well as non-linguistic functions, such as eye contact, body language and facial expressions".

Pragmatics (as defined above) is acknowledged as the most consistently and universally impaired linguistic domain in ASD (Asperger, 1944; Baron-Cohen, 1988; Dewey & Everard, 1974; Kanner, 1946; Tager-Flusberg, 1981; Tager-Flusberg et al., 2005; Young, et al., 2005), even in those children and adults who have had a history of ASD but no longer meet the criteria for a diagnosis on the spectrum (Kelley et al., 2006). Pragmatic deficits are, in fact, directly linked to the earliest social and communicative diagnostic criteria of the pathology.

Because pragmatics involves both verbal and nonverbal communication, profiles of abilities and weaknesses in this domain can be pinpointed regardless of the language level of the individuals. Many behaviours considered to fall under the umbrella of pragmatics are in fact nonverbal, social or affective in nature. In this sense, the existence of pragmatic deficits has been identified in all individuals with autism, including those who are minimally verbal or even nonverbal. Volkmar et al. (2005) provide the following list of the most commonly recognised nonverbal pragmatic deficits in ASD (3):

(3) a. Abnormal eye-gaze patterns;

b. Limited social referencing and sharing effect;

c. Inconsistent response to name;

d. Limited pretend play;

e. Limited or even absent social smiling;

f. Failure to understand and use conventional gestures;

g. Low frequency of facial expressions and nonverbal communication;

h. Limited interest in people and interaction;

i. Failure to point to objects of interest;

j. Low frequency of joint attention.

Shifting now to the verbal part of the autism spectrum, pragmatic difficulties associated with the use of language in contexts of social interaction have also been detected. The most commonly recognised atypical communicative behaviours are listed in 4 (a) to (d).

(4a) Echolalia, both immediate and delayed (Prizant & Duchan, 1981). Even if the function of echolalia is not well understood, (Eigsti et al., 2007) suggested that it “might serve several purposes, communicative and otherwise. Children might use echoing in conversation when unsure of their response; as a familiar verbal ritual; or as a way of holding information in memory” (see Chapter 5 for further discussion).

(4b) The use of jargon and neologisms, frequently reported as the invention and usage of nonsense terms and/or phrases with consistent meaning (Tager-Flusberg et al., 2005). It usually serves as a private communicative ‘‘bridge’’ between the child and his entourage (typically the parents or the family in general). However its function is currently still not completely understood.

(4c) Unique speech style, typically marked by flat or singsong intonation, lack of prosodic contour adapted to conversational expectations, suprasegmental speech qualities such as inappropriately soft, or, more frequently, loud speech volume, hoarseness, hyper-nasality, and unusually fast or slow speech rates (Shriberg et al., 2001).

(4d) Pronoun reversal, the erroneous use of deictic pronouns such as *you* for self-reference and *I* for an addressee, which has often been described as a core characteristic of language in ASD. The cause of this erroneous use of deixis may be due to basic impaired role-playing and perspective-taking skills (Evans & Demuth, 2012; Naigles et al., 2016; Novogrodsky, 2013).

From a more discourse-based perspective, the linguistic ability of people with ASD may be inappropriately adapted, characterized by a violation of turn-taking and the inability to adjust one’s conversational contribution with a marked reduction of cohesion and coherence in discourse (De Villiers et al., 2010). A typical example of impaired discourse abilities is the failure to respond adequately to questions and comments, with a consequent high rate of inappropriate responses (Capps et al., 1998).[3] Another well-known characteristic is the violation of Grice’s (1975) conversational maxims, including, in particular, saying things that manifestly lack relevance to the hearer, such as use of pronouns without a specific contextual referent (Baltaxe & D’Angiola, 1996). Spontaneous interaction with the hearer may be very limited, with a consistent lack of requests for information. When speech does occur, it often contains stereotypes and perseveration on a single topic, repetitive questioning (which are not requests for information) and affirmation by repetition. Another striking characteristic of discourse style in autism is weak mastery of prosody, both in expression and in comprehension, which seems to be widespread (Diehl et al., 2008). Finally, in comprehension, people with ASD may not project themselves into the speaker’s point of view and when speaking, they may not anticipate what the hearer will understand or want to know. Directly related to this, De Villiers et al. (2010) suggested that “one very

[3] We will further detail the topic of inappropriate responses in children with ASD in the paragraph dedicated to the choice of our experimental task and in the review of literature on morphosyntax.

important divide lies between those pragmatic competences which pertain to non-literal contents – as in, for instance, metaphor, irony and Gricean conversational implicatures – and those which pertain to the literal contents of speech acts". Many studies have indicated that individuals with ASD show extreme literalness: they have difficulties in understanding non-literal uses of language such as metaphor, irony and jokes and they display problems with the detection of hyperbole or understatement (Happé, 1994; Kaland et al., 2002; MacKay & Shaw, 2004; Norbury, 2005; Tager-Flusberg, 2000).

In sum, a large body of literature demonstrates that pragmatic impairment, despite phenotypical variability in severity and heterogeneity in linguistic capabilities, is the universal, life-long feature of language of the ASD population. The question then arises as to what the source of this pragmatic deficit may be. Following Eigsti et al. (2011), there are two main proposals in the literature: impaired ''Theory of Mind'' (ToM), and impaired "Executive Functions" (EF) theory. ToM theory suggests that difficulties in representing the contents of other people's minds are central to pragmatic impairment in ASD, and may provide a critical constraint on pragmatic language skills (Baron-Cohen, 1988). The false belief task and its nonverbal variants are one of the classic methods used in the study of ToM. Both "first-order", e.g. Sally-Anne task (Wimmer & Perner, 1983), and "second order", e.g. John and Mary task (Perner & Wimmer, 1985) false beliefs tasks have been widely accounted for being very difficult for most children and adults with ASD (Happe, 1995), while typically developing children tend to succeed at these tasks as of four years of age.

The second approach that tries to explain the core pragmatic and discourse deficits in ASD is the EF theory. The EF model suggests that ASD involves impairments in a set of cognitive processes associated primarily with the functional circuitry of the frontal lobes of the brain (Arnold et al., 2009; Ozonoff et al., 1994; Pennington & Ozonoff, 1996). These processes include: working memory, inhibition, set-shifting, goal maintenance, and cognitive control. The EF theory proposes that deficits in these processes may account for the symptoms of ASD, such as social deficits, communication delays, and repetitive behaviours. "By this account, children with autism may fail at pragmatic and discourse tasks because they are unable to simultaneously consider and respond to multiple sources of information (from self and others) or to inhibit inappropriate, potent, or salient responses" (Eigsti et al., 2011: 684). However even if this theory seems plausible, there is little specific evidence to support a specific role of EF in pragmatic abilities.

To sum up:

It is well established that pragmatic deficits are universal in individuals with ASD. Within the subgroup of verbal children, it has been argued that in some cases it is the only area of language that is deficient (i.e. individuals with Asperger's Syndrome). To date, however, neither the ToM theory nor the EF theory has been demonstrated to provide a full explanation for the symptoms of pragmatic impairment in ASD.

1.2.2 Structural language in ASD: some methodological considerations

Besides pragmatics, human communication involves other domains of linguistic knowledge: vocabulary (lexicon), syntax (grammar) and phonetic/phonology (sound structure). While the evaluation of language abilities in verbal children with ASD has typically been determined on the basis of vocabulary testing, formal aspects of language in ASD (phonology and morphosyntax) have received much less attention, especially in languages other than English. For very young children, language has generally been assessed indirectly through parent report instruments such as the MacArthur CDI Infant Form [Fenson et al., 2000] (Charman et al., 2003; Hudry et al., 2010) or the Children's Communication Checklist – CCC [Bishop, 1998] (Geurts & Embrechts, 2008) and through measures of cognitive functioning, such as the Mullen Scales of Early Learning, MSEL [Mullen, 1995] (Akshoomoff, 2006), a multi-domain measure of early development which includes scales for expressive and receptive vocabulary. For older children (and adults), receptive vocabulary tasks, such as the Peabody Picture Vocabulary Test (PPVT) [Dunn & Dunn, 1997], have frequently been used as broad indicators of general language abilities (McGregor et al., 2012). However, exclusive focus on lexical knowledge does not give an exhaustive picture of language capacities, as such knowledge is quite different in nature from the language abilities underlying structural language, notably phonology and morphosyntax. Moreover, it has been suggested that lexical ability involving knowledge of individual words can be a spared, or even enhanced, domain in children with autism (for an overview see Walenski et al., 2006; 2008), and that some tests, such as the Échelle de Vocabulaire en Images Peabody, EVIP (the French version of the Peabody, PPVT), may actually overestimate linguistic abilities, at least in children with High Functioning Autism, HFA (Mottron, 2004). These hypotheses are grounded on findings suggesting that lexical/semantic memory may serve, in individuals with autism, as compensatory mechanisms for aspects of procedural (and perhaps even episodic) memory (Mostofsky et al., 2000; Ullman, 2004), consistent with a relative sparing of lexical and semantic memory. However, to our knowledge except for one study (Sukenik, 2017) who looked at the relation between lexicon and morphosyntax in children with ASD, no one has specifically tested for, or reported on, enhancement in autism in lexical abilities in comparison with structural language abilities.

Moving to the combinatorial domains of phonology and morphosyntax, recent studies, targeting the description of language abilities in children with ASD, have started to investigate formal language abilities both in comprehension and production (see Eigsti & Schuh, 2017; Tuller at al. 2017 for overviews). One of the most striking outcomes of these works is the fact that results have been very heterogeneous and sometimes contradictory, yielding a kaleidoscopic picture of structural language abilities in children with ASD. In the following sections, our aim is to sort out this extremely diverse collection of results. This allowed us to target the methodological criteria we decided to adopt as a baseline for our research.

1.2.3 Terminology

Before moving to the current state of knowledge on formal language abilities in children with ASD, it is fundamental to specify the terminology that will be used. For the sake of clarity, we decided to separate children in two groups on the basis of their performance on tests evaluating structural language. If performance was in the normal range the children were referred to as ASD-LN (ASD and normal formal language abilities), and if performance was impaired, they were referred to as ASD-LI (ASD with formal language impairment). Why did we choose these labels? We discovered that after the study by Tager-Flusberg (2006), in which the author stated that there are essentially two language phenotypes in autism – ALI and ALN – reflecting subgroups in which structural language aspects are quite different, researchers started to generalize use of these two labels to refer, respectively, to children with ASD with and without formal language impairment. However these labels have been widely employed to indicate sometimes contrasting conclusions. Tager-Flusberg and colleagues introduced the ALI label with the specific intention of designating a subtype of ASD which is phenotypically identical to SLI. Other authors (Loucas et al., 2008 a.o.), instead used the ALI/ALN labels for merely identifying two subgroups of language abilities, concluding in fact against the idea that there is a subtype of children with SLI among children with autism. This contradictory use of the ALI label has created a great deal of confusion in the literature. In an effort to avoid this confusion, we chose to adopt the terms ASD-LI and consequently ASD-LN as more neutral, descriptive terms. They should be interpreted as a description of the performance of children with autism on formal language abilities, excluding any *a priori* relation with SLI. We will return to this topic in section 1.3.

Concerning cognitive abilities we decided to apply to the totality of the studies reported in Chapter 1 a generalisation of the two most frequent labels used for defining cognitive abilities. In our review, we will refer to children with High Functioning cognitive capabilities as “HF” and children with Low Functioning cognitive capabilities as “LF”. Why? First of all, the global labels referring to cognitive abilities vary considerably from one study to another. Across the different studies, participants have been divided into profiles of cognitive abilities (HF and LF) via psychometric evaluation of intelligence, through the use of a wide variety of standardized tests and indices (indices for general cognitive abilities, fluid reasoning, developmental scales, nonverbal IQ, etc.) and through the use of different cut-offs (the cut-off for cognitive impairment varied between standard scores of 70 to 85). Moreover, picturing the heterogeneity of the ratio behind the characterisation of profiles of cognitive abilities would have created further confusion in our review. We will return, in detail, to the methodological problem of evaluating cognitive abilities in children with ASD in the second chapter of the present work. Finally, the same group categorisation was applied to the few studies including toddlers with ASD. Normally toddlers with autism are not assessed via psychometric batteries with full scale IQ scores, meaning that cross-sectional literature studies are necessarily skewed. To overcome this limitation, in studies including toddlers with ASD, the HF/LF distinction was made on the basis of their mental age (if their mental age corresponded to their chronological age we considered them as HF).

1.2.4 Articulation and Phonology in ASD

In the literature that has addressed the question of phonological development in children with ASD, findings have been somewhat conflicting. Some studies have indicated, across wide age ranges that phonology seems to be the area of language least likely to be impaired in people with autism, while others have identified mild or even severe phonological impairment. However, it should be mentioned that studies have not systematically distinguished between articulation (speech) and phonology, very often collapsing these two terms in the broad term "*phonological abilities"*. We propose, first of all, to give clear definitions for articulation and phonology, in order to allow for clear interpretation of the results of the relevant studies that we will report on in the following section.

Articulation refers to the movement of the speech organs (tongue, lips, larynx, teeth, hard palate, velum, jaw, nose, and mouth) to produce sounds. It overlaps with phonetics, which refers to the physical production and articulation of sounds. Disorders that impact the form of speech sounds are traditionally referred to as articulation disorders and are associated with structural (e.g., cleft palate) and motor-based difficulties (e.g., apraxia).

Phonology is the study of how sounds are organized and combined sequentially and hierarchically to form syllables and words in natural languages. Speech sound disorders that impact the way speech sounds (phonemes) function within a language are traditionally referred to as phonological disorders. They result from impairment in the phonological representation of speech sounds and speech segments—the system that generates and uses phonemes and phoneme rules and patterns within the context of spoken language. The process of perceiving and manipulating speech sounds is essential for developing these phonological representations.

In the following paragraphs we report on current knowledge explicitly differentiating articulation and phonology. We identify which abilities (articulatory or phonological) were evaluated and what kind of error analysis was adopted in each study (articulatory or phonological errors). These precautions were fundamental because several studies have conducted analyses of phonological abilities through the use of tasks that were designed for the evaluation of articulatory abilities.

The next section will be organised as follows: first we will report on articulatory abilities and then on phonological abilities. Each part will be divided between studies that have investigated HF children with ASD and studies which have included both HF and LF children with autism. Results showed that, while it has been claimed that articulation and phonology are a relatively spared domain in children with ASD (especially when compared to other linguistic domains such as morphosyntax), several studies have identified mild or even severe articulatory/phonological impairment. We will conclude our review with two studies which compared phonological impairment found in children with ASD-LI and in children with SLI, and we will introduce the first study which investigated phonological abilities in autism through the use of a nonword repetition task (Kjelgaard & Tager-Flusberg, 2001). The introduction of such a task will be indicated as a turning point in the literature on phonological abilities in autism. We will come back on this topic in detail in section 1.2.10. A summary of all of the studies that are explored in the next sections, and their over-all results in terms of whether evidence was provided for impaired phonology (ASD-LI) or spared phonology (ASD-LN) is provided in **Figure 1** at the end of the section.

1.2.4.1 Articulatory and phonological abilities in children with ASD

Articulatory (or speech) processing in ASD has been sparsely investigated and has been reported to be either age appropriate or superior compared to other expressive language abilities (Rapin & Dunn, 2003). In one of the first studies conducted on a sample of children with autism, Bartak et al. (1975) found that 47 HF children with ASD (mean age 7;0) had very few articulatory problems on both a structured and a spontaneous task of speech production. Similar findings were reported by Boucher (1976), who, using the Edinburgh Articulation Test (Anthony et al., 1971), found superior articulatory abilities in 31 HF children with autism (mean age 11;2 years old) in comparison to an age-matched TD control group and a younger group of children with receptive “dysphasia” (SLI). In contrast, Shriberg et al. (2001) in a study of conversational speech data involving 30 individuals with HFA/AS aged 10-50 years old found a high prevalence of residual articulatory and speech errors in the groups with pathology (almost 80% for both the HFA and the AS), most frequently expressed as distortions of specific sounds, such as sibilant dentalization and lateralization, deviant articulations or errors in combining consonants and vowels. The authors concluded that a deficit in such processes could underlie residual articulatory errors that could persist for a lifetime in the ASD population. More recently, Cleland et al. (2010) evaluated 30 children with HFA and 39 children with AS (the mean age of both groups was 9;6) with a standardized test of articulation (GFTA-2, Goldman & Fristoe, 2000). The GFTA-2 tests the accuracy of 39 different English phonological elements (single consonants and consonantal clusters) in isolated words (word initial, medial and final positions). Although the authors applied this test with the specific intent of evaluating phonological abilities in children with autism, they ended up in describing their articulatory capacities. Results showed that a subgroup (12%) of participants presented standard scores below norms, indicating a speech disorder. Although all the other children had standard scores within the normal range, a sizeable proportion (33% of those with normal standard scores) presented a small number of errors. Overall 41% of the HFA group produced at least some speech errors. The errors of the children with ASD were mainly characterized by processes observed in typical development (in particular, gliding, cluster reduction and final consonant deletion); however non-developmental articulatory error types (such as phoneme specific nasal emission, e.g. /s/ and /z/, and initial consonant deletion) were also found both in children performing below the normal range and in those who performed within the normal range. Interestingly, previous studies on adolescents and adults with ASD reported similar non-developmental errors, suggesting that this impairment does not improve over time.In addition, some researchers have identified a small number of children with HFA having ‘‘extraordinary difficulty producing intelligible speech’’ (Catherine Lord & Paul, 1997). One very well-known case is that reported by Wolk & Edwards (1993), who described language abilities in an 8-year-old boy with autism whose speech was so severely disordered that he was unintelligible. This child showed an HF profile with both typical and atypical patterns of articulatory development. Children with severe articulation disorders are often excluded from research because of the difficulty of assessing their articulatory level via standardized tests. Consequently, very little is known about the phenomenology of their language impairment. To our knowledge, no study has investigated articulatory abilities in children with LF cognitive abilities.

A greater number of studies have examined **phonological** abilities in autism. Evidence for unimpaired phonological abilities in children with ASD has been found, both in reception and in production. Norbury et al. (2010), for example, conducted an eye-tracking experiment with thirteen HF children aged 6;7 to 7;9 and thirteen age-matched TD children. The main aim of their study was to determine whether information, semantic or phonological, is encoded in the initial form-referent mappings of new words. Children were evaluated longitudinally on two trials. The first one addressed the learning of new words, while the second served as a consolidation trial. Children with ASD were more successful than their TD peers at mapping phonological forms to novel referents. However, as the authors pointed out, the ASD advantage was not maintained over time. At the second trial session, TD children showed clear consolidation of learning both semantic and phonological information, while children with ASD did not. Concerning investigation of expressive phonological abilities, Schoen et al. (2011) analysed the production of 30 toddlers with ASD, aged 19-36 months, in clinician–child structured play sessions. They compared the production of the ASD group with that of two TD control groups, one age-matched and the other language-matched. Results showed that children with ASD had phonological systems that functioned like those of children at similar levels of both age and language development, in terms of percentage of correct consonant production in words. Analysis of syllable complexity presented a similar picture, with the ASD group reaching rates of correct production similar to those of language matched controls, and only slightly lower from those of more linguistically advanced age-mates. The authors concluded that "toddlers with ASD have phonological systems that function much like those of children at similar levels of language development, when broad phonemic categories are analysed" (2011: 185), indicating that in a developmental trajectory, phonology in ASD is only slightly delayed (and not deviant).

Other studies have found significant phonological deficits in HF children with autism, indicating that phonological impairment can be present despite a normal IQ. Bartolucci & Pierce (1977) tested three groups of children (10 HF children with ASD aged 10 years old, 10 "mentally retarded" aged-matched children and 10 TD mental age-matched children) for phonological abilities, both in production and in reception (in particular, stops, fricatives, affricatives, nasals and liquids). Production was assessed through a picture naming task, designed to elicit production of the 24 consonant phonemes of English in initial, final, preconsonantal, postconsonantal, and intervocalic positions. Reception was assessed through a speech discrimination task including the same stimuli as in the production task. An analysis was run on the interaction of phonological error patterns between the groups. Results indicated that the autistic group had phonological impairment similar to that demonstrated in the subjects with intellectual disability. Moreover a comparison of the sum of the percentages of errors in production and in perception demonstrated a predominance of perception errors in all groups.

As far as we know, only four studies have analysed the phonological capabilities of children presenting deficits in the cognitive domain. Kjelgaard & Tager-Flusberg (2001) conducted a study on 89 children with LF and HF autism (FSIQ range 25-141) aged 7;4 years old in average. The children were administrated a battery of standard language tests, targeting their phonological, lexical and higher-order semantic and grammatical abilities (the Clinical Evaluation of Language Fundamentals, CELF-R (Semel, Wiig, & Secord, 1987)). They were given a test of articulation measuring the accuracy of productive phonology, the GFTA (Goldman & Fristoe, 1986), and a test of repetition of Nonsense Words, the NEPSY subtest (Korkman et al., 2007). Results on the GFTA indicated that phonological skills were almost always spared in autism overall (more than 90% of the sample), despite severe deficits in vocabulary, and in semantic and grammatical knowledge. Moreover, scores on the GFTA test were not correlated with FSIQ, indicating that even LF children with autism were able to perform normally on this kind of task. In contrast, the authors discovered that the participants had much more trouble with the nonsense word repetition test. This finding may seem somewhat surprising given what was observed on the GFTA test and because of the general belief about phonological capabilities in ASD at the time. This study opened up a new perspective on the possible existence of subtle phonological impairments in ASD. We will discuss it further detail in section 1.2.10.1.

McCleery et al. (2006) compared 14 children with LF autism (mean age 3;4 years old) with 10 language-matched TD children, on an imitation task of consonants production. They collected speech samples over several sessions of interaction between each child and an experimenter, who was instructed to prompt different types of English consonants. Overall results showed that the ASD group level on phonological development generally followed a normal TD-like trajectory.

The third study which included LF children was conducted by Rapin et al. (2009) on a cohort of 62 children with ASD with both HF and LF profiles and a mean age of 8;6 years old. They evaluated both syntax (through a composite score of language comprehension) and phonology (production only). For the latter, they used the Photo Articulation Test (PAT) (Lippke et al., 1997), which yields a score for correct speech sounds produced in naming photographs and objects. The task, purported to assess phonological abilities, consists in naming multiple series of pictures and thereby inducing production of a consonant or a vowel in different positions (e.g. /p/ in initial “pie”, medial “apple” and final position “cup”). Each production is scored as correct or incorrect (sound substitution, omission or distortion). Results supported the existence of two major types of phonological profiles in school-aged children with autism: the first type was represented by 76% of the children with ASD, who presented intact phonological abilities and above average cognitive abilities. The second type, representing 24% of the sample, was characterized by severely impaired expressive phonological skills at age 8;6. However, within this second profile, the authors emphasized the existence of two clusters of children: cluster 1 (17.7% of the group) which consisted of children who presented low nonverbal cognitive abilities associated with profoundly impaired phonology and syntax; and cluster 2 (only 6% of the sample) which consisted of children who differed drastically from the first cluster in having average NVIQ’s and average syntactic abilities, in the face of profoundly impaired phonological expression. The authors had not clinically detected this kind of selective phonological impairment profile before (Rapin & Dunn, 1997). They gave it the label of “verbal dyspraxia”, arguing that neurologists had detected oromotor deficits in both clusters of the impaired phonological group. However, we would argue that it is not possible to conclude clearly whether the impairment found in the study is due to an articulatory deficit or to a phonological impairment, due to the fact that the PAT test is constructed to evaluate articulatory abilities, and that the final score of the task mixes articulatory and phonological abilities.

Finally, Tuller et al. (2017) examined phonological abilities in a group of 20 children with ASD, both LF and HF, aged 6- to 12-years old, via the word repetition task of the Bilan Informatisé de Langage Oral, BILO battery (Khomsi et al., 2007) and they compared their performance with 20 age-matched children with SLI. In this task children are asked to repeat real words, selected according to two criteria: syllable length and phonological complexity. Each item was coded for identical repetition, the number of errors produced, together with the syllable position and error type (i.e. substitution, addition, omission, etc.). Results showed that a subgroup of children with ASD (*n* = 14) had particular difficulty with consonants in syllable final position, notably liquids and obstruents in internal coda position (e.g. /ɔʁlɔʒ/ 'clock' and /tʁaktoeʁ/ 'tractor'). Impaired performance of these children resembled that found in the group of children with SLI. Both groups of language-impaired children were found to use the same strategies in order to avoid syllabic complexity, such as omission of a segment (e.g. /sɔti/ instead of /sɔʁti/ 'exit'). Moreover, the nonverbal cognitive abilities of these children were not correlated with performance on phonology. Among the children, be they with LF or HF autism, some displayed phonological impairment and some did not.

1.2.5 General conclusion on articulatory and phonological abilities in children with ASD

Taking into account the preceding review of studies reporting on articulatory and phonological capabilities via tasks of reception and production of words and/or analysis of spontaneous speech samples, we can draw the following conclusions:

1. The investigation of **articulatory** abilities in children with ASD has been limited to HF autism. The current *status quaestionis* on articulatory abilities in the ASD population is fundamentally conflicting in two main directions:

    a. Individuals with HF autism present essentially spared articulatory capabilities (Bartak et al., 1975; Boucher, 1976; Cleland et al., 2010).

b. A subgroup of individuals with HF autism present impaired articulatory capabilities which are deviant from the normal developmental trajectory and persistent with age (Cleland et al., 2010; Lord & Paul, 1997; Shriberg et al., 2001; Wolk & Edwards, 1993).

2. The current *status quaestionis* on **phonological** abilities in the ASD population can be divided between studies that have investigated performance in children with HF autism and studies that have included both HF and LF children with autism:

   a. A subgroup of children with HF autism present essentially spared phonological abilities, similar to those shown by TD age peers (Kjelgaard & Tager-Flusberg, 2001; Norbury et al., 2010; Rapin et al., 2009; Schoen et al. 2011; Tuller et al., 2017).
   b. A subgroup of children with HF autism present impaired phonological abilities (Bartolucci & Pierce, 1977; Kjelgaard et al., 2001; Rapin et al., 2009; Tuller et al., 2017).
   c. A subgroup of children with LF autism present spared phonological abilities, (Kjelgaard & Tager-Flusberg, 2001; McCleery et al., 2006; Tuller et al. 2017).
   d. A subgroup of children with LF autism present impaired phonological abilities, like those found in children with SLI (Kjelgaard & Tager-Flusberg, 2001; Rapin et al., 2009; Tuller et al. 2017).

To sum up:

Two profiles of articulatory and cognitive abilities have been detected in the literature:

1) ASD-LN with HF cognitive abilities
2) ASD-LI with HF cognitive abilities

No conclusion can be drawn on articulatory abilities in children with LF abilities, since no study has looked at this population.

Four profiles of phonological and cognitive abilities have been found in the literature:

1) ASD-LN with HF cognitive abilities
2) ASD-LI with HF cognitive abilities
3) ASD-LN with LF cognitive abilities

4) ASD-LI with LF cognitive abilities

Two studies have compared the performance of children with ASD-LI and children with SLI and they both found striking similarities between the two groups.

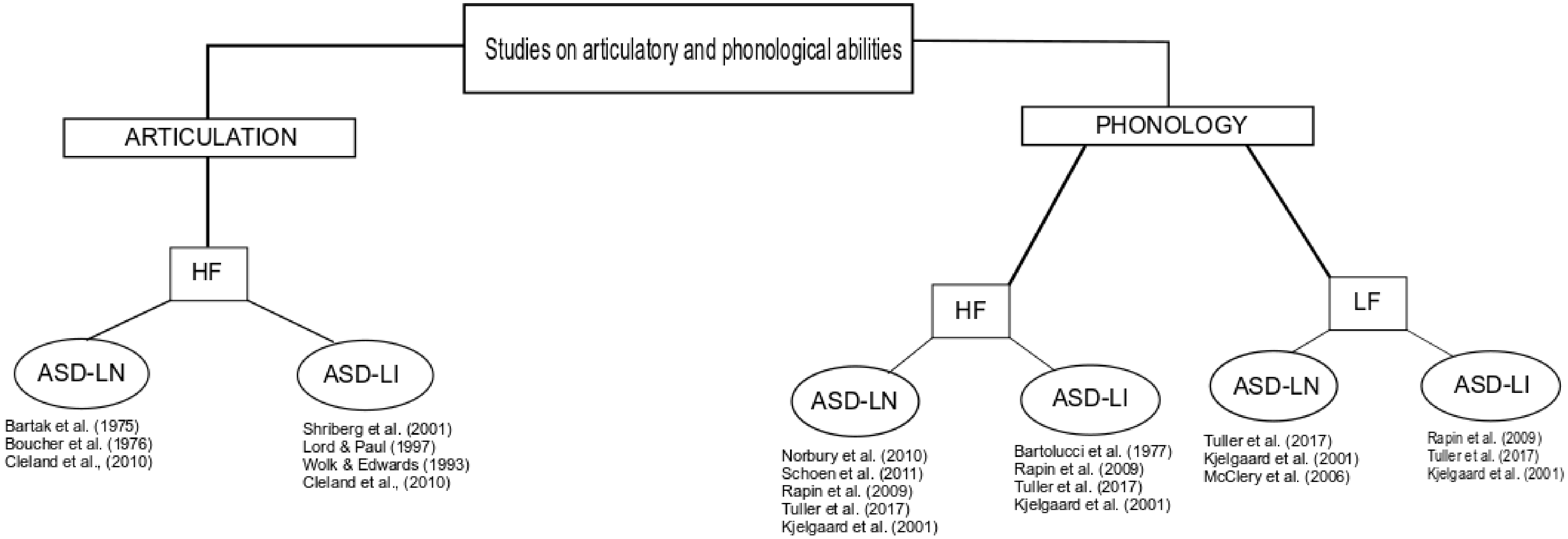
In conclusion: Until now, the debate on the existence of a phonological impairment in children with ASD has addressed the following topics: whether children with ASD in general have spared or impaired articulatory abilities, whether they have spared or impaired phonological abilities and whether their articulatory/phonological capabilities are related to their cognitive (dis)abilities. The great heterogeneity of results indicates that much more work should be done on the characterisation of phonological abilities in children with ASD. Notably the introduction of phonological evaluation via the use of nonword repetition tasks may cast new light on the phonological abilities of children with ASD. Kjelgaard and Tager-Flusberg, comparing the same children on both a task of word-repetition and a task of nonword repetition found that children with ASD performed much worse on this last one. We suggest that performance on tasks of articulation, repetition of real words and spontaneous speech samples may be the result not only of phonological knowledge, but of several other factors. In particular, tasks of articulation and repetition of real words may engage the use of pre-existing lexical knowledge, which in relation to word familiarity or frequency effect can lead to biased performance (Coady & Evans, 2008). We will report in detail about studies that have used task of nonword repetition for evaluating phonological abilities in further detail in section 1.2.10.

Summarizing, we have highlighted the following limitations in the studies of phonological abilities in children with ASD published so far:

I. Standardized tasks and analysis of spontaneous speech samples may not be sufficient to describe the phonological abilities of children with ASD. The variability in reported results from might be a consequence of the heterogeneity of both the assessment measures and the tools.
II. There is a lack of clarity about the difference between articulatory abilities and phonological abilities. The main purpose of our study was to focus on phonological capacities incorporating an evaluation of phonological complexity.
III. There is a lack of studies including both HF and LF children and a comparison with children with SLI. Both cognitive profiles should be investigated in order to verify whether phonological abilities are related to intelligence level.

IV. No study has analysed the possible relation between severity of autism symptoms and phonological abilities in children with ASD. We need to verify that poor performance on phonological tasks is not due to severity of autism symptom rather than to actual shortcomings in language abilities.

FIGURE 1. SUMMARY OVERVIEW OF STUDIES EVALUATING ARTICULATORY AND PHONOLOGICAL ABILITIES IN CHILDREN WITH ASD



1.2.6 Morphosyntax in ASD

Syntax is the set of rules, principles, and processes that govern the structure of sentences in any given natural language. The term morphosyntax encompasses both syntactic and inflectional morphology rules to represent syntactic features via morphological marking according to their occurrence in particular syntactic contexts. Morphosyntactic development in children with ASD has come under increasing scrutiny in the last decade, yielding conflicting findings. A large number of studies have suggested that morphosyntax seems to be a relatively spared domain (especially when children's mental age is taken into account), while other studies have identified mild or even severe morphosyntactic impairment. The existence of an impaired profile, has led some researchers to formulate the hypothesis that there is indeed an SLI profile in ASD. This parallel has been drawn, in the first instance, for children with HF cognitive abilities, since SLI is a neurodevelopmental disorder that involves specific difficulties in mastering structural aspects of language, independently from any kind of intellectual disability. Although some studies have found very striking similarities between ASD-LI and SLI in general performance on morphosyntactic tasks, the same studies have highlighted different error patterns, questioning the actual resemblance of the two conditions. Crucially all of these studies have left open the question of how language abilities develop in children with LF abilities. Few studies on language in ASD have included children with both normal and impaired cognitive abilities. Results have been conflicting, with some studies having found that LF cognitive abilities generally entailed impaired language and a few studies evoking the existence of a profile with normal language abilities. In other words, some studies have raised the question of whether the similarities found between ASD-LI and SLI in the HF population are extendable to verbal children with LF abilities and language impairment. The results of the comparison ASD-LI/SLI have yielded very similar conclusions to those found for HF children.

The next sections will report on current knowledge based on the literature about morphosyntactic abilities in children with ASD. We organised our literature review by separating those studies that investigated exclusively children with HF autism from those studies that included both HF and LF children with autism. This division should ensure a clearer picture of the current work that has been done on morphosyntactic abilities in this population and it will highlight the imbalance between the number of studies having looked only at children with HF autism and those having included children from the whole spectrum. A summary of all of the studies explored is reported in **Figure 2**, at the end of the section.

1.2.6.1 Morphosyntactic abilities of children with HF autism

In the following paragraphs we will review studies that have investigated morphosyntactic abilities (via comprehension and production tasks) in children with HF autism. We will refer first to studies that compared morphosyntactic abilities of children with ASD with those displayed by TD children. The structure of the section will be the following: first we will report on studies that investigated general morphosyntactic abilities in HF children with ASD (grammatical morphemes, combining words to form complex sentences, MLU, etc.); then we will present studies that investigated specific morphosyntactic constructions in HF children with ASD (grammatical marking of tense and agreement, wh-questions, relative sentences, passive sentences, clitics and pronouns, and binding). We will conclude that studies can be divided between those which identified a group of children presenting normal formal language abilities (ASD-LN, autism with normal language) and those which also identified a group presenting impaired formal language abilities (ASD-LI, autism with language impairment).

Since most studies investigating children with ASD-LI have suggested that the shortcomings displayed by these children resembled the ones showed by children with SLI, we will then report on comparative studies between these two impaired populations. After a brief report on the theoretical debate about the possible *continnum* or co-morbidity of these two conditions and their developmental trajectories, we will concentrate on those studies that directly investigated morphosyntactic abilities in ASD and SLI. We will concentrate first on studies that compared HF autism to SLI, returning, in the next section, to those that included children with ASD-LI with both HF and LF autism. We will conclude that, since children with ASD-LI and children with SLI have been reported to have similar global performance but different error patterns, we cannot draw clear conclusions on the actual nature of the similarities between children with ASD-LI and children with SLI.

1.2.6.1.1 Studies comparing morphosyntactic abilities in children with HF autism and TD children

The investigation of morphosyntactic abilities in children with ASD has been brought to the forefront in recent years. Although early work often reported the development of morphosyntax to be TD-like, there is currently general agreement among scholars regarding the existence of mixed performance in children with ASD: some children behave like TD children, but others do not.

Looking at studies that investigated general morphosyntactic abilities in children with HF autism, a pioneer cross-sectional study, (Waterhouse & Fein, 1982), found that the development of grammatical morphemes (Brown's 14 grammatical morphemes of English, which included: present progressive –ing, the use of proposition *in* and *on*, regular plural –s, irregular past, possessive, uncontractible copula, articles, regular past –ed, regular third person singular of present tense –s, irregular third person singular (*does*, *has*), uncontractible auxiliary, contractible copula and contractible auxiliary) in 33 children with ASD, aged 5-15 years old, without any cognitive disability showed the same order of acquisition as in TD children. However, more recently, Tek et al. (2014) conducted individual growth curve analyses on the same expressive morphosyntactic measures (Brown's 14 morphemes) as well as wh-questions and sentence complexity (MLU) using a longitudinal dataset of spontaneous speech of 17 children with HFA (mean age = 32 months). Results showed the existence of two different profiles: a language profile that displayed grammatical development highly similar to that of TD children (ASD-LN), and a profile of children that showed an expressive language delay reminiscent of a global impairment in expressive language (ASD-LI), coupled with impairments in other areas of development, including high rates of autism severity.

These discrepancies between the studies suggested that children with ASD may show different levels of abilities on different morphosyntactic constructions. This state of affairs has led some researchers to develop experimental tasks which allowed for investigation of specific aspects of morphosyntax. Looking at the grammatical marking of tense and agreement on verbs, Naigles et al. (2011) reported that HF pre-school-aged children with ASD (aged 27 to 37 months) can show TD-like grammatical use of plural markers, past tense and mapping of novel verbs into transitive frames in English via an experimental task of syntactic bootstrapping. Similarly, a longitudinal study by this same lab, based on a preferential looking paradigm, indicated that pre-schoolers with HF autism understood the inflectional marker for the present progressive (-*ing*) at the same level as younger TD children (Goodwin et al., 2012; Tovar et al., 2015). However, other studies have found consistent impairment in the processing of the same constructions. Eigsti & Bennetto (2009) evaluated a group of HF children and adolescents (aged 10 to 16 years old) with autism and a group of TD children on the grammaticality judgment task used by (Johnson & Newport, 1989) in their well-known study on second language acquisition. Results indicated a significant group effect (ASD < TD) for 3rd person singular marking and present progressive marking. Interestingly, grammatically judgments appeared to be related to severity of autism symptoms in the ASD group. Zhou et al. (2015) showed that Mandarin-speaking children with autism, aged 4 to 6 years old, also exhibited deficits in using grammatical morphemes that mark aspect.

Concerning wh-questions, Goodwin et al. (2012) found that pre-schoolers with HF autism understood these constructions at the same level as TD children. In the same perspective Jyotishi et al. (2017) concluded from a longitudinal study of off-line comprehension of wh-questions that pre-school children with ASD (mean age = 36 months) showed a developmental pattern similar to the one displayed by TD children. Similarly, Su et al. (2014) showed that HF children with ASD (mean age 6;6 years old) were relatively strong in understanding the linguistic properties specific to the interpretation of Mandarin wh-questions.

Similar results of unimpaired syntactic abilities were found by Schaeffer (2017), who explored the syntactic comprehension and production of object relatives in Dutch-speaking HF children with ASD, aged 5-13 years, reporting that they did not differ from TD age-matched children.

Two studies from Terzi and colleagues found impaired performance both in comprehension and production of clitic pronouns in Greek children with HFA (aged 5 to 8) compared to TD vocabulary-matched children. The same results were found both when children were evaluated on comprehension via an experimental picture matching task (Terzi et al., 2014) and on production via the narrative task *Frog, where are you?* (Terzi et al., 2017).

Regarding c-command and principles of Binding theory, studies have found that these principles/properties are generally spared. Diehl et al. (2015) reported that HF children and teenagers with ASD (7-17 years old) were able to efficiently use prosodic cues to resolve sentences involving syntactic ambiguities at the same rate as TD peers. Similarly, Janke & Perovic (2015) and Khetrapal & Thornton (2017) showed that knowledge of c-command and Principle A were intact in children with HFA in comparison with age-matched controls, respectively via a picture-matching task and a grammatical judgment task (on groups of 7;3-16;4 year-old participants). However, some of the children with HFA and spared performance on the Binding theory performed below norms on a standardized test of morphosyntactic comprehension (TROG-2 Bishop, 2003), indicating impaired language abilities.

To sum up:

Reports of good grammatical performance have been widely found for HF school-aged children and teenagers with ASD, when compared to the TD population. Nevertheless, a high number of studies seem to contradict the claim that children with autism generally have good or even intact use and development of grammatical abilities, even those who are high functioning:

1. Some children with HF autism display essentially spared morphosyntactic abilities: (Diehl et al., 2015; Goodwin et al., 2012; Janke & Perovic, 2015; Jyotishi et al., 2017; Khetrapal & Thornton, 2017; Naigles et al., 2011; Schaeffer, 2017; Su et al., 2014; Tovar et al., 2015; Waterhouse & Fein, 1982).
2. Some children with HF autism display impaired morphosyntactic abilities: (Eigsti & Bennetto, 2009; Eigsti et al.,, 2007; Ellis Weismer et al., 2011; Kjelgaard & Tager-Flusberg, 2001; Park et al., 2012; Tek et al. 2014; Terzi et al. 2014; 2017; Zhou et al., 2015).

We can conclude that: in HF children with ASD there are (at least) two profiles of language abilities; these two profiles can be identified as ASD-HF with normal language abilities (ASD-LN) and ASD-HF with impaired language abilities (ASD-LI).

1.2.6.1.2 Studies comparing morphosyntactic abilities in children with ASD-LI and children with SLI

One of the issues raised in research on morphosyntactic abilities in children with ASD is to what extent the performance of children with ASD-LI is comparable with the characteristics of language of children with SLI. This question has been accompanied by a theoretical debate about the *continnum* or the *phenomimicry* of these two conditions. The next section will report on current knowledge on the topic. After summarizing briefly the theoretical debate about the similarities between ASD-LI and SLI, we will move to the studies that have directly compared morphosyntactic performance in these two conditions. After the literature review, we will conclude that the similarities between these two conditions are not so clearly defined when the morphosyntactic abilities of children with ASD-LI and children with SLI are compared via qualitative error analysis.

1.2.6.1.3 Structural language in ASD: is there an overlap with SLI?

In the last decades there has been long-lasting interest in the relationship between SLI and ASD. According to conventional diagnostic frameworks, SLI and ASD are mutually exclusive diagnoses. Nevertheless, diagnostic frameworks do not necessarily reflect clinical reality, and over many years the idea has been explored that there might be overlapping language deficits in the two conditions (Bishop, 2010). Lately, the diagnostic dissociation between SLI and ASD has been questioned, causing a flourishing debate between the possible *continuum* and *phenomimicry* of these two conditions (Bishop, 2010; Tomblin, 2011; Williams et al., 2008). Although a deficit in structural language is not core to autism, individuals with ASD can manifest, independently of diminished NVIQ, clinically significant impairment in structural aspects of language that resemble those found in children with SLI, as first attested in the milestone study by Bartak et al, (1975). In more recent years, Kjelgaard & Tager-Flusberg (2001) showed, via a series of standardized language tests commonly used to diagnose SLI, that 76% of their 89 children with ASD performed in the impaired range. Similarly, Tager-Flusberg (2006) found that in a group of 35 children with ASD 57% displayed a profile that closely resembled the one reported for children with SLI. Finally, in a population screening sample (SNAP), Loucas et al. (2008) found that 57% of their 97 children with autism and normal NVIQ had impaired performance on a language battery. Demonstrating that the occurrence of LI in ASD seems to be consistently high, Tager-Flusberg and colleagues proposed that this relationship may consist of a partial overlap between SLI and ASD, suggesting that the two disorders were points of a continuum of severity. Therefore, according to this approach, there is a subgroup of children with ASD that display both SLI and ASD. In this framework the diagnostic overlap is referred to as "*continuum*", referring to the two diseases as intimately connected. One interpretation of this hypothesis is to regard autism as 'SLI plus', i.e. to assume that the only factor differentiating the disorders is the presence of additional impairment in autism. Tager-Flusberg and colleagues have widely supported the existence of a common language profile, called ALI, not only via an investigation of formal language abilities but also on the basis of genetic studies and brain structure abnormalities in both SLI and ALI (Leyfer et al., 2008; Lindgren et al., 2009; Tager-Flusberg, 2015; Tager-Flusberg & Joseph, 2002).

Other researchers, notably Bishop (2003, 2006, 2010), Conti-Ramsden et al., (2006), Whitehouse et al., (2007), and Williams et al., (2008), have expressed scepticism over the hypothesis of a continuum between the two pathologies. Although they generally agree that children with ASD-LI and children with SLI are globally similar in their formal linguistic performance, they suggest that a qualitative examination of the data can demonstrate that the similarities between the two deficits are actually superficial and present different underlying causes at the cognitive, neurobiological, and/or etiologic levels. These studies have also indicated that structural and pragmatic language deficits, although logically separable, often co-occur in children with ASD. This vision of the relationship between the two pathologies has been described as "*phenomimicry*", since it considers the two disorders distinct. Bishop (2010: 623) defines this model as *phenomimicry* "In the case of ASD and SLI, this would mean that having a causal trait for one disorder could lead to a clinical picture resembling the other disorder. […] One version of phenomimicry [may be] a child with ASD who develops language deficits similar to those in SLI. The converse is also possible: a child with SLI might develop a clinical picture resembling ASD, perhaps because social interaction is difficult and stressful." In sum, in a phenomimicry model, children with ASD-LI would present ASD symptoms similar to those seen in children with ASD but without language impairment and language abilities similar to children with SLI. As suggested by Tomblin (2011), it may be more reasonable to conclude, in this perspective, that there is a large group of children with ASD who have poor structural language skills, rather than a group of children with both ASD and SLI.

No matter which of these two views is adopted, the question remains: why are there so many children with ASD who have poor structural language abilities? Besides looking deeply into neurobiological and etiological levels of explanation, we think that we should try to explore more carefully the linguistic phenotype in children with ASD and language impairment, in order to see whether they indeed show the exact same profile as children with SLI. Following the recent trend in the literature concerning the investigation of formal language abilities in ASD, we will report on the current knowledge on this topic.

1.2.6.1.4 ASD-LI and SLI in HF autism

Recent studies have concentrated on very specific syntactic constructions and they have shed light on the presence of specific morphosyntactic impairment not clearly attributable to general cognitive abilities in children with ASD-LI, as is the case for SLI. It has been claimed that these two conditions appear to share, for some individuals, analogous shortcomings in general morphosyntactic abilities. These similarities have been investigated via standardized tests, such as the CELF-III by Lloyd et al. (2006) in 18 children with SLI, 10 HF children with ASD and 9 HF children with ASD-LI all aged between 5;10 and 10;7, and by Loucas et al. (2008) in a cohort of 41 HF children with ALI, 31 HF children with ALN and 25 children with SLI; and via the CELF-4 by Ellis Weismer et al. (2017) in three groups of 21 children with SLI, 27 HF children with ASD and 32 TD children 9;6 years old. While Ellis Weismer et al. (2017) have detected a similar profile between the children with ASD-LI and children with SLI, with children performing with very similar rates in both groups, the other two studies have found mixed results. Lloyd et al., (2006) have found that language impairment in ASD-LI can be less severe than it is SLI, while Loucas et al., (2008) have detected a more severe impairment in ASD-LI. For these three studies, however, the "Core Language Score" of the CELF battery was used to compare children with ASD-LI and children with SLI. No details on subtests performance were provided and, since the CELF is not conducive to evaluation of performance on specific grammatical structures, it was very hard to interpret results.

Other studies have examined morphosyntactic abilities in both ASD-LI and SLI groups via experimental tasks targeting specific morphosyntactic constructions. Tager-Flusberg, (2006) assessed a group of 35 HF children with ASD aged between 7 and 14 years old (*M* = 10;4), using conversational speech samples taped during parent-child play interactions and examiner-child interactions which were part of the administration of the ADOS. Results showed that, when compared to children with a group of age-matched SLI (*n* = 13), children with ASD-LI did not differ in production of grammatical marking of tense, such as 3rd person *–s* and past *–ed*. Sukenik & Friedmann (2018) looked at complex constructions, e.g. relative clauses and wh-questions in Hebrew in a group of 18 HF children with ASD aged 9;0-18;0 (*M* = 13;4) via tasks of picture comprehension and elicited production. The authors found that two subgroups of abilities were present in the ASD group: a subgroup of children displaying normal language and a subgroup with language impairment (ASD-LI). Depending on the analysis used to evaluate morphosyntactic abilities, children with ASD-LI performed in line with children with SLI on global performance, while from a qualitative error analysis different behaviours between children with ASD-LI and children with SLI emerged, notably a more severe and pervasive language impairment in children with ASD-LI. The authors concluded that simply looking at general scores does not make it possible to draw firm conclusions about the actual resemblance between ASD-LI and SLI profiles.

To sum up:

In the previous section we have concluded that investigation of morphosyntactic abilities of children with HF autism seem to indicate the existence of two main profiles: the first one involves normal development (ASD-LN) and the second one moderate to severe impairment (ASD-LI). This second profile has been compared to the one that characterizes children with SLI. Some studies have directly compared the performance on morphosyntactic tasks of children with ASD-LI and children with SLI. Results have been contradictory:

1. Some studies found that children with ASD-LI and children with SLI displayed same performance (Ellis Weismer et al. 2017; Tager-Flusberg, 2006).
2. Some studies found that the language impairment in children with ASD-LI can be less severe than the one found in children with SLI (Lloyd et al., 2006).
3. Some studies found that children with ASD-LI are more impaired than children with SLI (Loucas et al., 2008; Sukenik & Friedmann, 2018).

In conclusion: Some of the studies reported here found that the performance of children with ASD-LI resembled the one of children SLI. One study highlighted differences in error patterns, questioning the actual resemblance of the two conditions. Comparison made on general scores is not sufficient to describe similarities and differences that can occur between children with ASD-LI and children with SLI. We will see in the next section how some researchers have included a qualitative analysis of errors in order to better describe possible phenotypical overlaps between these two populations.

1.2.6.2 Morphosyntactic abilities of children with LF autism

Research exploring morphosyntactic abilities in children with LF cognitive abilities has been comparatively sparse. The reason for this lack of research stems from the fact that the equation "LF abilities = low (or even absent) linguistic abilities" has been taken for granted. However, in very recent years, some studies have started questioning this *a priori*, and have begun including verbal children with LF cognitive abilities in their research samples.

In the first part of the next section, we will report on the few studies (seven) that have included both HF and LF children in their investigation of morphosyntactic performance. Results will show that, although all of these studies have found that the vast majority of children with LFA show impaired formal language abilities, two studies also found a subgroup of children with spared morphosyntactic abilities.We will conclude for the existence of two profiles of language abilities, ASD-LN and ASD-LI, in children with LF cognitive abilities, just as has been found for HF autism.

Moreover, although the parallelism with SLI was designed in first instance to compare children with HF abilities and ASD-LI, some studies raised the question of whether the similarities found between these two populations were extendable to verbal children with LF abilities and language impairment. In the second part of the next section, we will report on studies that compared children with ASD-LI (both HF and LF) with children with SLI. Through this literature review we will try to answer two questions: 1) whether performance of children with ASD-LI (which in this case includes both HF and LF children) on morphosyntactic tests differed from that of children with SLI. In order to answer this question we will report on both the quantitative and qualitative analyses run in the studies and 2) whether the occurrence of LF cognitive abilities in ASD-LI would further lower the linguistic performance of these children or whether their performance is no different from that of children with ASD-LI and HF abilities. Concerning the first question, results will show that while children with ASD-LI (both HF and LF) may share the same general performance with children with SLI, a qualitative analysis of their performance highlights some striking differences. We will therefore speculate on the possible nature of these discrepancies between general performance and qualitative error analysis, reporting the hypothesis of some researchers that the observed ASD-SLI differences could largely be due to pragmatic deficits in ASD, rather than to a qualitative difference in structural language skills. Concerning the second question, we will see that a large majority of the studies did not find a relation between language and cognitive impairment, suggesting that impaired intellectual abilities may not have an impact on linguistic performance in verbal children with ASD. Finally we will report on the very few studies (two) that detected lower morphosyntactic performance in children with ASD-LN when compared with TD age-matched children.

1.2.6.2.1 Morphosyntactic abilities in children with both HF and LF cognitive abilities

Few studies have investigated morphosyntactic abilities in children presenting ASD and intellectual impairment. Ellis Weismer et al. (2011) showed that impairment of syntactic abilities in both children with HF and LF autism can be detected starting at very early stages. When matched on overall vocabulary level and nonverbal cognition, 30-month-old toddlers with autism exhibited early grammatical abilitieswhich were qualitatively very similar to 25-month-old late-talking toddlers without autism. evaluated grammatical complexity in the language of children with ASD as measured through the Index of Productive Syntax (IPSyn), a tool which calculates syntactic complexity in speech by evaluating the occurrence of 56 syntactic and morphological structures (relative clauses, wh-questions, complex infinitives, etc.). They found that, even when adjusting for nonverbal mental age, levels of syntactic complexity were reduced in children with ASD. Similarly, (Park et al., 2012) evaluated 17 children with HF and LF autism (no number was given for each group) on some subscores of the IPSyn. Results showed that, for a subgroup of these children, the use of verb phrases noun phrases and sentence structures were areas of impaired development, while the other skills were intact (questions and negation). Language performance was not related to the cognitive level of the participants. (Perovic et al., 2013) investigated knowledge of *Principle A of the Binding Theory* in 18 children with HF and LF autism (no number was given for each group). Results provided support for the proposal that Principle A is either missing or impaired in children with ASD-LI and low cognitive abilities, in comparison to children with ASD-LI and high cognitive capacities. Durrleman et al., (2017) assessed (in addition to providing measures of standardized language tests) the comprehension of passives in 20 French-speaking children with ASD aged 7;8 to 10;11, both HF (n = 11) and LF (n = 9). In this study the authors showed that not only did the children with ASD-LI show problems in comprehending passive sentenced but that a significant qualitative differences was also present between the children with ASD-LN and the TD children on morphosyntactic abilities. A recent study by Jensen de López et al. (2018), investigated the comprehension of passive sentences in a group of 15 Danish HF (*n* = 13) and LF (*n* =2) children with ASD. Results showed that the children with HF abilities did not have problems in comprehending long and short periphrastic passive sentences. The authors reported that the only two children who showed language impairment were the ones displaying a deficit in cognitive capacities.

As can be seen, although some studies found a relationship between low intellectual abilities and impaired language, children with ASD and LF abilities have not been found to systematically display impaired performance. The existence of this other “discrepant” profile, ASD-LN with low IQ, has also been evoked in other studies. Kjelgaard & Tager Flusberg (2001) evaluated a group of 89 school-aged children with ASD via the CELF-III (Semel, Wiig, & Secord, 1995). The total sample was composed of both children with HF cognitive abilities and children with LF cognitive abilities, however there was no clear indication of how many children were intellectually impaired. Results showed that the HF children were divided between those who displayed normal language abilities and those with impaired language. Children with low IQ generally displayed more difficulties in completing tasks evaluating structural language. In fact, 45 out of 89 children were not able to perform the CELF in its entirety and they were subsequently excluded from the analysis. Crucially, this group was composed primarily of LF children. Amongst the 44 children that completed the CELF, however, some had LF cognitive capabilities. Although the vast majority of these LF children displayed impaired performance on the CELF, a few of them (around 10% of the total sample, percentage we calculated on the basis of the few indications given by the authors) were able to perform in the normal range. Joseph et al. (2002) evaluated language abilities in a population of 120 children with either HF or LF autism through use of a parent-reported score for phrase speech production (from the ADI-R), which is defined as the spontaneous, flexible use of at least two words in combination, one of which must be a verb. Similarly to Kjelgaard & Tager-Flusberg (2001), this study found two subgroups of language abilities among children with HF cognitive profile (ASD-LN and ASD-LI) and two subgroups of language abilities among children with LF cognitive profile, one with ASD-LI (which was composed by the majority of the LF children) and another with ASD-LN. This second profile constituted roughly 8% (10/120 children) of the sample.

To sum up:

Reports of impaired grammatical performance were found in all eight studies that investigated morphosyntactic language abilities in children with LF autism. Two of these studies, however, also reported the existence of a subgroup of LF children that displayed normal language abilities:

1. Both children with HF and LF autism seem to display impaired language abilities (Durrleman et al., 2017; Eigsti et al., 2007; Ellis Weismer et al., 2011; Park et al., 2012)
2. Children with LF autism seem to display essentially impaired language abilities (Jensen de López et al., 2018; Kjelgaard & Tager-Flusberg, 2001; Perovic et al., 2013).
3. A subgroup of children with LF abilities seems to display spared morphosyntactic performance (Joseph et al., 2002; Kjelagaard & Tager-Flusberg, 2001).

We conclude that: in LF children with ASD there are (at least) two profiles of language abilities; these two profiles can be identified as ASD-LI with LF abilities and ASD-LN with HF abilities. This second profile has been detected in studies with a high number of participants and in very small number of children, raising the question of its relative prevalence in the ASD population.

1.2.6.2.2 Evaluation of morphosyntactic performance in studies including both children with HF and LF autism and children with SLI

Several recent studies have compared morphosyntactic abilities in children with ASD to those of children with SLI, but these studies have mostly focused on HF children. This has left open the question of language abilities for the rest of the spectrum (LF), which, we racall, constitutes almost a third of the ASD population. To our knowledge only six studies have run a comparison of morphosyntactic abilities between children with SLI and children with ASD with either spared or impairedintellectual abilities. All of these studies have targeted, via experimental tasks, specific morphosyntactic constructions.

Modyanova et al., (2017) and Roberts et al., (2004), investigated elicited production of grammatical marking of tense trough picture description, such as 3rd person *–s* and past *–ed* English. Neither of these two studies directly included children with SLI in their protocol. However, results of children with ASD were compared with performance of children with SLI, based on what is found in the literature. Modyanova and colleagues' study included 83 children with autism (aged 4;3 to 16 years old) with HF and LF profiles (although they did not indicate how many children displayed an intellectual impairment). Roberts et al. assessed 62 children with ASD aged 5 to 15 years old, among which 43 with HF or borderline cognitive abilities and 19 with LF autism. Four studies directly compared children with ASD and children with SLI, including both populations in their sample.

Durrleman & Delage, (2016), Tuller et al., (2017) and Prévost et al., (2018) investigated the performance of children with ASD on pronominal clitics in French. The first study included 21 individuals with ASD aged 5-16 as well as 22 individuals with SLI also aged 5-16. Among the children with ASD, nine presented a LF profile. The second and the third study, assessed production of pronominal clitics in 20 participants with ASD aged 6 to 12 and 20 age-matched children with SLI. Four participants with ASD showed impaired intellectual abilities. Finally two studies investigated comprehension and production of wh-questions in French (Prévost et al., 2017; Zebib et al., 2013). Participants in both studies were the same as in Tuller et al. (2017).

Looking at general performance (quantitative analysis), only one study found that language impairment in ASD-LI can be more severe than it is in SLI (Modyanova et al., 2017), due probably to severity of autism symptoms, according to the authors. The other studies detected a similar profile between the two groups on a number of different structures (Durrleman & Delage, 2016; Roberts et al. 2004; Prévost et al., 2017; Tuller et al., 2017; Zebib et al., 2013). However, qualitative error analysis revealed that, alongside the similarities in structural language performance displayed by the ASD-LI and the SLI groups, children with ASD produced inappropriate answers more frequently than the children with SLI. Notably, Roberts et al. (2004) reported that the children with ASD far exceeded the comparative group of children with SLI in the production of non-target answers. Similar results were found by Zebib et al.'s (2013) on production of wh-questions and by Prévost et al. (2017) in both production and comprehension of wh-questions. For example, in Prévost et al., (2017), children with ASD were elicited for the production of wh-questions. In this case, the examiner showed the child a picture containing a character performing an action and a hidden part. The child was then told to ask the character a question about the hidden part (e.g., "Look, here the rabbit is pushing someone, but we can't see who. To know who the rabbit is pushing, ask him"). Following the child's response, the hidden part was then revealed. A recurrent inappropriate production of children with ASD was to give an answer to the question, guessing who was under the hidden part (e.g; children response 'a coconut'). Moreover, children with ASD were more affected by perseveration strategies in their answering pattern. Durrleman & Delage (2016) found very similar performance on object clitic production in their children with ASD-LI and their children with SLI, with the only exception of one substructure (object clitic 1st person), where the children with SLI outperformed those with ASD. The authors appealed to a deficit in theory of mind abilities in the children with ASD to explain their lower performance on 1st person clitics. However, when the same phenomenon was tested in the study by Tuller et al. (2017) no difference was reported either on general performance or on the types of errors produced. Prévost et al. (2018) compared the two studies and argued that task differences could be the cause for diverging results. Some children in the ASD-LI group used non-target answers quite often, but some children with SLI used such answers as well, and in comparable proportions. Spontaneous production was also investigated by Tuller et al. (2017), who highlighted a particular difficulty of children with ASD-LI in producing constructions involving deep clausal embedding, compared to children with SLI. They advanced the hypothesis that this low embedding rate in the children with ASD may be due to their difficulties sustaining a conversation due to communication impairment.

Even if these two groups seem to perform at the same impaired rate in morphosyntactic tasks, the reasons underlying their low performance could be different. The existing studies, reviewed here, make it difficult to conclude that the performance of the children with ASD-LI is similar to children with SLI. Some authors (Prévost et al., 2017; Zebib et al., 2013) proposed that pragmatics, may have an effect on performance of children with ASD on tasks designed to assess structural language, and may partially obscure fundamental similarities and/or differences in the structural language difficulties in the two populations. These difficulties also raise the complementary question of the extent to which low structural language performance could be the result of pragmatic difficulties, in (some) children with ASD. We will explore this question in the next section.

Regarding the issue of children with LF abilities among the ASD-LI group, the impaired performance of these children has not been systematically reported to be related to their level of intelligence. Modyanova et al. (2017) was the only study that reported a strong relation between structural language impairment and low cognitive abilities; the other studies did not find such a correlation (Durrleman & Delage, 2016; Robert et al. 2004; Prévost et al., 2017; Tuller et al., 2017; Zebib et al., 2013). These results seem to suggest that the linguistic performance of children with ASD-LI is not lowered by impaired cognitive abilities. One very interesting result that corroborates this hypothesis was represented by the identification of one child (out of six) with LF autism and spared language abilities in the study of Tuller et al. (2017), which is reminiscent of the profile found in the studies by Kjegaard & Tager-Flusberg (2001) and Joseph et al. (2002).

Finally, it is interesting to notice that two studies detected qualitative differences between children with ASD-LN and TD children on morphosyntactic abilities. Modyanova et al. (2017) and Tuller et al. (2017) found that when tested on specific complex constructions (grammatical marking of tense and clitic pronouns), children with ASD-LN performed slightly lower (albeit in the normal range) than age-matched controls, suggesting that despite a very good phenotypical realization of language skills, the language of ASD-LN children may not be entirely normal. These results raise new questions on the language abilities of children with ASD-LN.

To sum up:

Very few studies have taken up the question of morphosyntactic performance in both children with HF or LF autism and children with SLI. Results showed that despite similarly low performance rates in children with ASD-LI (either HF or LF) and children with SLI, the two groups differed in the kinds of errors that they made.

1. One study found that independently of their cognitive level (HF or LF), children with ASD-LI performed significantly lower than children with SLI (Modyanova et al., 2017).
2. Several studies found that children with ASD-LI (either HF or LF) and children with SLI displayed same general (low) performance (Durrleman & Delage, 2016; Robert et al. 2004; Prévost et al., 2017; Tuller et al., 2017; Zebib et al., 2013). Once qualitative error analysis was taken into account, however, these studies found that children with ASD-LI (either HF or LF) behaved differently from children with SLI. One possible explanation for these discrepancies is that pragmatic deficits may have an effect on performance on tasks designed to assess structural language in ASD, and may partially obscure fundamental similarities or differences in the structural language difficulties in the two populations.
3. Except for one study (Modyanova et al., 2017), no relation between low cognitive abilities and language impairment has been found. Nonetheless, all these studies had fewer LF children than HF children, which makes it hard to make firm conclusions on the nature of the relationship between intellectual and morphosyntactic abilities.
4. A few studies found that based on a quantitative and qualitative error analysis of morphosyntactic performance, ASD-LN children had lower scores than TD age-matched children (Modyanova et al., 2017; Tuller et al. 2017).The same result was found also in Terzi et al., (2017) and Durrleman et al. (2017), who argued for

differences of performance between children with AD-LN and age-matched TD peers.

In conclusion: no clear conclusions can be drawn on the actual similarities and differences between the performance of children with ASD-LI (either HF or LF) and children with SLI. The low structural language performance of children with ASD-LI could be further aggravated by the well-known pragmatic deficit associated with ASD, which is reflected by a tendency to produce inappropriate responses when the task requires use of language in context. Moreover the fact that some studies highlighted some qualitative differences between ASD-LN profile and TD children called into question the real nature of the normal language abilities of these children. Disentangling these factors could be explored through studies based on tasks that focus more directly on measures of structural language, thereby reducing the possible effects of pragmatic impairment in ASD-LN profile.

1.2.7 General conclusions on morphosyntactic abilities in children with ASD

Considering the studies that investigated morphosyntactic abilities in HF and LF children with ASD we can draw the following conclusion:

Four profiles of morphosyntactic and cognitive abilities have been found:

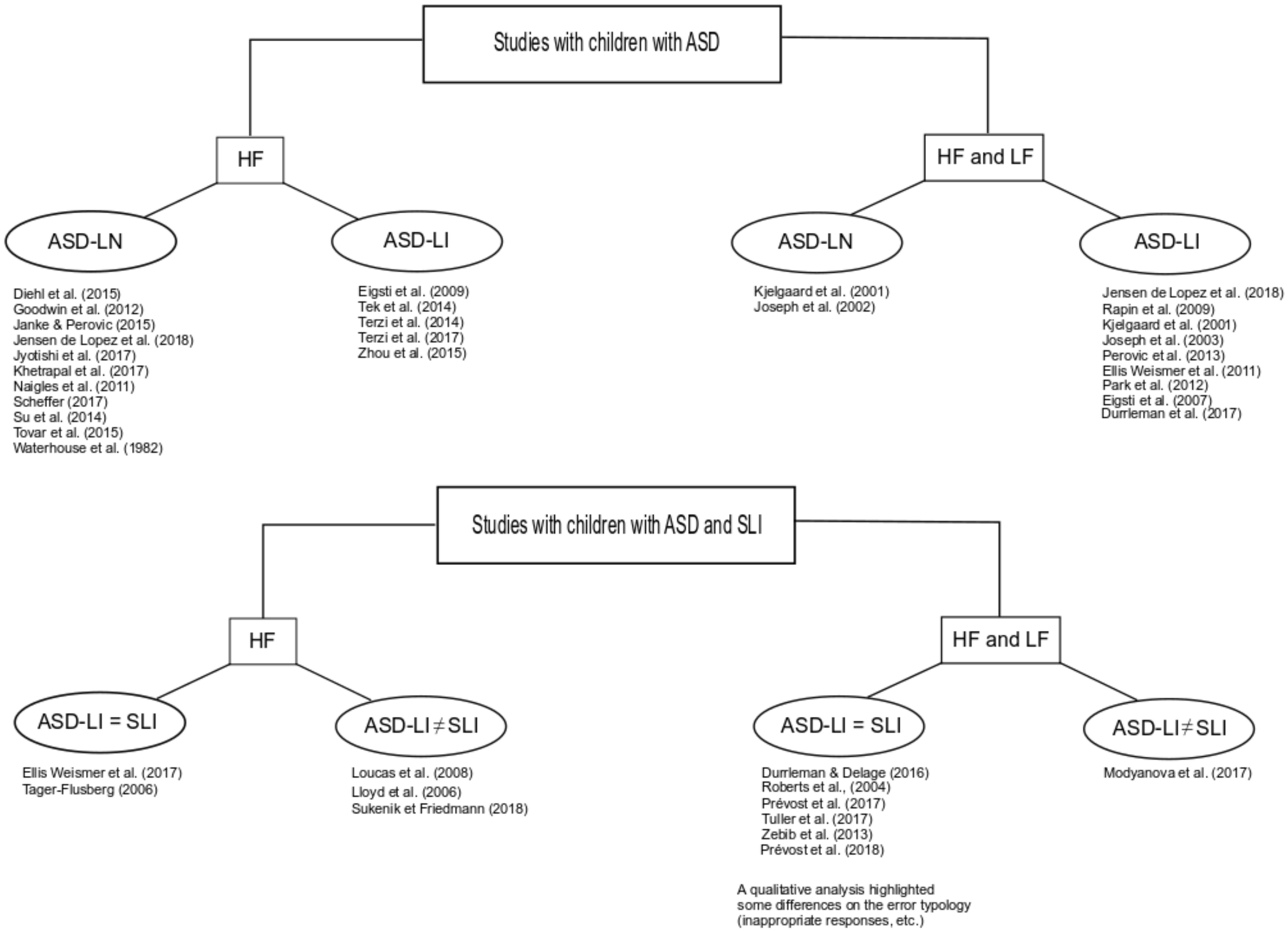
1) ASD-LN with HF cognitive abilities
2) ASD-LI with HF cognitive abilities
3) ASD-LN with LF cognitive abilities
4) ASD-LI with LF cognitive abilities

The profile of children with ASD-LI, independently from cognitive abilities (either HF or LF), has been compared to the one found in children with SLI, while the profile of children with ASD-LN has been compared to TD children. Results showed that on quantitative performance the ASD-LI and SLI profiles look similar and the ASD-LN profile resembles the TD profile. However, when compared on the basis of qualitative error analysis, striking differences come to light for each comparison. It is still not clear whether these differences are due to the choice of measures and tools for assessing morphosyntactic abilities or to genuine differences between ASD-LI and SLI. In order to answer this question we need to detect and reduce the confounding effects that some factors may create in children with ASD. In other words we should focus our attention on those types of tasks that can reduce the effects of communicative deficits in this population. The aim is not to ignore pragmatic impairment but to minimize its effect on evaluation of other areas of language, in order to obtain "purer" measure for structural aspects of language. These considerations have been explicitly expressed by some researchers in their choice to use repetition tasks to evaluate morphosynatctic abilities in verbal children with ASD. We will expand on the current literature on this topic in section 1.2.11.

1.2.8 Final considerations on structural language abilities in children with ASD

In conclusion to this section we can say that very heterogeneous and often contrasting results have been found for formal language abilities of children with ASD, especially on the relationship between ASD-LI and SLI. We believe that these results justify our decision to refer to these subgroups as ASD-LN and ASD-LI profiles instead of ALI and ALN in our literature review. There is currently no clear reason to conclude that these two conditions are part of a *continuum*. Moreover, we could not derive a clear picture of the relationship between cognitive abilities and linguistic abilities. For now, we can conclude that both in phonology and morphosyntax all four logically possible profiles of structural language/cognitive abilities have been found.

FIGURE 2. SUMMARY OVERVIEW OF STUDIES EVALUATING MORPHOSYNTACTIC ABILITIES IN CHILDREN WITH ASD



1.2.9 Exploring language phenotypes in ASD via the use of repetition tasks

In the preceding literature review, we saw that recent studies targeting the description of language abilities in children with ASD have raised the possibility that neither available standardized tests nor many experimental tasks targeting specific aspects of structural language may be sufficient to isolate the source of impairment in children with ASD (Wittke et al., 2017 a.o.). In particular, measures drawn from standardized tests rarely enable detailed analysis of specific phonological and morphosyntactic structures. These tests generally target multiple aspects of several language domains such as morphology, lexicon, syntax, phonology and articulation, at the same time, preventing distinct evaluation of these abilities. This state of affairs has led some researchers to develop experimental tasks which allow for investigation of specific aspects of morphosyntax, as we saw in the previous sections. However, these tasks do not appear to be devoid of cofounding effects either. They frequently require functional communication skills and competence in social interaction (e.g. act-out, sentence-picture matching, elicited tasks), which are known to be universally impaired in individuals with ASD. In particular, very recent studies have pointed out that low performance of children with autism on language tests targeting morphosyntax may be the result of a misunderstanding of the pragmatics of the testing situation (Prévost et al., 2017; Tuller et al., 2017; Zebib et al., 2013). Children with ASD have been shown to have a tendency to produce high rates of inappropriate answers (no responses, perseveration strategies) on different kinds of tasks, both in production and comprehension. Prévost et al. (2017) suggested that difficulties in appropriate use of language in context (pragmatics) could have an effect on performance on tasks designed to assess structural language, and may ultimately obscure our understanding of the difficulties that children may have with structural aspects of language, including fundamental similarities with SLI. In turn, performance on tasks of articulation and repetition of real words may be a result not only of phonological knowledge, but several other factors. Tasks of articulation and repetition of real words may engage the use of a pre-existing lexical knowledge, which in relation to word familiarity or frequency effect can create biased performance (Coady & Evans 2008).

In parallel, research on SLI has recently focused on the identification of “clinical marker tasks” or “endophenotypes” of language impairment (following (Moll et al., 2015), such as Nonword Repetition (NWR) and Sentence Repetition (SR). Both standardized repetition tasks (e.g. ‘CNRep’ by Gathercole et colleagues and ‘recalling sentences’ of the CELF battery) and experimental tasks of NWR and SR have been pinpointed as good sources of information about children's phonological and morphosyntactic-levels in clinical assessment. These tools have yielded higher levels of sensitivity and specificity, in comparison with standardized tests evaluating morphosyntactic and phonological abilities, among individuals with SLI across a great number of languages, such as English (Coady & Evans, 2008; Conti-Ramsden et al., 2001; Redmond et al., 2011; Riches et al., 2010; Seeff-Gabriel et al., 2010), Cantonese (Stokes et al., 2006), Cypriot Greek (Theodorou et al., 2017), Czech (Smolík & Vávru, 2014), Dutch (Rispens & Parigger, 2010), Swedish (Kalnak et al., 2014), Slovak (Kapalková et al., 2013), Turkish (TopbaŞ et al., 2014), Italian (Bortolini et al., 2006; Devescovi & Caselli, 2007; Dispaldro et al., 2013), and French (Leclercq et al., 2014; Thordardottir et al., 2011).

Only a few studies have addressed the diagnostic accuracy of repetition tasks in children with ASD. The hypothesis is that since these tasks are very sensitive to SLI, they should also be able to detect structural language impairment among children with ASD. Moreover, if the phenotypical realization of formal language impairment in children with ASD corresponds to the one displayed in SLI, the two groups should perform alike on these tasks, both at the quantitative and qualitative levels. Children with ASD would then be divided between those who show normal language abilities and those who manifest formal language impairment similar to what is found in children with SLI. Pragmatic impairment typical of children with ASD should have limited impact on repetition accuracy in SR tasks (Polišenská et al., 2015; Williams et al., 2013). The design of SR incorporates formal aspects of morphosyntactic processing but is less constrained by pragmatic features. Differently from other tasks evaluating morphosyntax, the child is not asked to create a sentence with a given word (formulating sentences), finish a sentence on the basis of a picture, answer an item-based question (elicited production), take the perspective of another person (act-out) or engage in a dialogue producing a story (story telling), etc. In SR, children are instructed to repeat sentences exactly as heard without any given pragmatic context. Analogously, the design of a NWR task should ensure a more controlled evaluation of phonological abilities, since individuals cannot rely on their pre-existing lexical knowledge, when repeating nonwords conceived to be unrelated to existing words in the child's language.

In the next paragraphs we will report what has been learned about formal language abilities (phonology and morphosyntax) in children with ASD through study of their performance on NWR and SR tasks. Now that we have framed the theoretical debate on ASD and SLI we can explicitly refer to the original labels of language abilities used in the different studies (ASD-LN/ASD-LI, ALN/ALI, etc.), with the aim of placing each study in its specific theoretical framework. After reporting on the main results found in the literature, we will describe the strengths and limitations of studies that have investigated structural language abilities thorough the use of NWR and SR tasks. Then we will argue in favour of the use of repetition tasks, motivating the introduction of two particular such tasks, the Language Impairment Testing in a Multilingual Setting, LITMUS Sentence Repetition task (LITMUS-SR) and the Nonword repetition task (LITMUS-NWR), developed for detecting SLI in monolingual and bilingual children, within COST Action IS0804, in both research and clinical practice.[4]

1.2.10 Nonword Repetition (NWR) Tasks

As seen at the end of the section 1.2.4.1, Kjelgaard & Tager-Flusberg was the first study to draw a parallelism between the performance of children with ASD-LI and children with SLI on a NWR task. This tool is known to be very sensitive to the detection of SLI among the general population (Conti-Ramsden et al., 2001). Moreover, since individuals cannot rely on their pre-existing lexical knowledge, NWR tasks were originally used to avoid any word familiarity or frequency effects that could affect the repetition of real words (Coady & Evans, 2008). However, not all NWR tasks are designed in the same way. Depending on its structure, NWR can be a relatively pure index of phonological short-term memory (Archibald & Gathercole, 2007 a.o.) or a strong discriminator of phonological complexity (Kjelgaard & Tager-Flusberg, 2001 a.o.).

In the first case, the longer the nonword, the stronger is the effect of working memory on repetition performance. It is well known that this length effect is dramatically present in most children with SLI and it starts to appear when nonwords have more than two syllables. That is why in most nonword repetition tasks evaluating phonological short-term memory in SLI and TD populations, nonwords can have up to five syllables or sometimes even more (Poncelet & Van der Linden, 2003). The presence of long nonwods, it is argued, can help discriminate between children with SLI and TD-age peers. Importantly, the syllabic structure of the nonwords included in these tasks almost systematically involved recursive CV syllables, and did not display internal complexity factors (we will come back on this

[4] COST Action IS0804 (2009-2013) developed a series of tasks for Language impairment testing in multilingual settings (LITMUS); see www.bi-sli.org.

topic in section 4.5.1.2). In the second case, the NWR task design has been conceived to concentrate, through the use of complex phonological structures, on those structures that could be identified as the source of errors in children with impaired phonology, independently from memory capacities of individuals. These kinds of NWR tasks limit the effect of phonological short-term memory (the nonwords usually don't excess 3-4 syllables) and word likeness (to limit the effect of lexical knowledge), in order to focus on syllable structure, which seems to be a suitable marker for assessing phonological disorders (e.g. complex consonant clusters).

When tasks of nonword repetition have been employed in children with ASD, phonological impairment has been consistently found, even when the presence of a phonological impairment did not emerge on other phonological tasks, as showed in Kjelgaard & Tager-Flusberg (2001). However, the conclusions that can be drawn from impaired performance vary according to the design of the NWR task. The hypothesis of researchers that have used the NWR tasks tapping phonological short-term memory in children with ASD, is that if a subgroup of children, ASD-LI, share the same performance as children with SLI, it is because both groups display impaired performance directly correlated to the increasing length of nonwords. On the other hand, the hypothesis of researchers that have used NWR tasks built on phonological complexity, is that if a subgroup of children, ASD-LI, share the same aetiology as SLI, it is because both groups show impaired performance directly correlated to the complexity of the syllabic structure and not only on nonword length. "Because [a nonsense word repetition task] is posited to involve both disassembling the input into smaller units and then reassembling these units in production, it is expected to involve compositional processes. The evidence therefore suggests that at least compositional aspects of phonology may be somewhat impaired in ASD" (Kjelgaard & Tager-Flusberg, 2001).

The vast majority of the studies that have compared children with ASD to TD age-peers and/or children with ASD-LI to children with SLI, have employed phonological short-term memory NWR tasks. Only a couple of studies have used NWR tasks built on phonological complexity. Once again, results have highlighted that the ASD-LI/SLI parallelism can be inferred only on a general level. Children with ASD-LI have been found to be in fact much less affected by syllable length than children with SLI. In the next section we will outline the main findings of the studies that have employed an evaluation of phonological abilities via a NWR task. At the end of the section we will highlight the strengths and the limitations of these studies and we will argue for the use of the LITMUS-NWR task for a purer evaluation of phonological abilities in children with ASD.

1.2.10.1 Literature review on studies that employed NWR tasks

As far as we know eleven studies have assessed phonological abilities in children with ASD via nonword repetition tasks.

Kjelgaard & Tager-Flusberg (2001): as said above, the NEPSY subtest Repetition of Nonsense Words was administered to a group of 89 children with LF or HF autism (IQ range 25-141) and a mean age of 7;4 years. This phonological short-term memory test is built on thirteen nonsense words of increasing length (3 two-syllable nonwords, 3 three-syllable nonwords, 4 four-syllable nonwords and 3 five-syllable nonwords). Participants were awarded 1 point for each syllable repeated correctly. The first striking result was the fact that only 40 participants (out of 89) were able to perform the task (cf. 79/89 on the GFTA). Among them, 31 had an IQ within the normal range (> 80). Among those who completed the repetition test, 34/40 had a score lower than one standard deviation below the mean. The authors concluded that this phenotype of phonological impairment was similar to what is found in SLI, although no control group of children with SLI was included in the study. The study did not provide any individual results, particularly for 9/40 children with an IQ below the normal range that were able to complete the task. Moreover, we do not know if the children that were able to succeed on the GFTA test were the same ones that completed the nonword repetition task. However, a few years later (Tager-Flusberg, 2006), the authors stated in reference to their previous work that "although there was a moderate relationship between language scores and IQ, there were children with high and low IQ scores, and normal or delayed onset of language milestones, in both the ALN and ALI language subgroups".

Botting & Conti-Ramsden (2003): they investigated phonological abilities in a group of 13 HF (IQ > 70) children with ASD (*M* = 10;10) and two groups of age-matched children, children with SLI and children with Pragmatic Language Impairment (PLI), via a task of verbal/phonological short-term memory consisting of 40 nonwords CNrep (Gathercole et al., 1994), divided equally into two-, three-, four- and five-syllable items. Half of the nonwords contain consonant clusters and the remainder have single consonants only. Results showed that children with SLI were the ones who performed most poorly on this measure, followed by peers with autism and children with PLI.

Bishop et al. (2004): a test of nonword repetition (Baddeley et al., 2004) was administrated to 80 children diagnosed with autism or PDD-NOS (*M* = 10;44, *SD* = 4.51). The test was composed of 36 items of increasing length (2-5 syllables). Children were cognitively assessed via WISC-III and had average IQ (*M* = 82.68, *SD* = 25.31). They were compared to a group of 59 age-matched TD children. Results showed that the children with ASD performed significantly lower than the controls on phonological processing in the nonword repetition task. However, there was no indication of disproportionate phonological difficulties in the children with ASD and normal verbal abilities. When the authors divided the children into two subgroups on the basis of their verbal IQ, only the children with a verbal IQ < 77 showed consistent impairment in the nonword repetition task (eleven children were unable to complete the task; all of them had a verbal IQ below 65). The authors concluded that "whatever factor leads to low nonword repetition in a subset of individuals with autism appears to be the same factor that causes low verbal IQ" (Bishop et al., 2004: 57).

Whitehouse et al. (2008): the same task used by Kjelagaard & Tager-Fluberg (2001), the NEPSY nonword repetition task, was administrated to a group of 34 children with SLI (*M* =11;10, *SD* = 2.3) and to a group of 34 HF (IQ > 80) children with autism (*M* =10;11, *SD* = 3.1) further divided in a group with poor structural language sills called Apoor, (*n* = 18) and a group with appropriate structural language skills called Aapp, (*n* = 16) on the basis of the results on a battery of standardized language tests. Results showed that the children in the Apoor group performed better than the individuals with SLI when the stimuli were four syllables or higher in length. In other words, the Apoor group was relatively less affected by stimulus length than individuals with SLI. This difference led the authors to conclude that the mechanism underlying poor nonword repetition in the Apoor group may be different from that underlying impaired nonword repetition in SLI. Since further analysis found that children with autism who also had poor nonword repetition had also substantial deficits in two or more autistic domains, the authors indicated severity of autism symptoms as a probable cause of poor linguistic abilities in children with ASD (see Whitehouse et al., 2007).

Riches et al. (2011): they explored nonword repetition skills in 16 adolescents with ALI (*M* = 14;8, *SD* = 5.77) and 13 adolescents with SLI (*M*= 15;4, *SD* =7.26), through the CNRep, (Gathercole & Baddeley, 1996) and the Nonword Memory Test (NMT, Gathercole & Baddeley,1989). The authors clearly stated that the purpose of combining the two assessments was to increase the number of stimuli, and henceforth the power of the study.

The stimuli were designed to be phonotactically similar to real English words, and therefore contain complex onsets and nuclei, e.g. consonant clusters, long vowels and diphthongs. Some of the words contained derivational morphemes, e.g. *-ing* and *-ate*. In total, children had to repeat 17 two-syllable words, 17 three-syllable words, 15 four-syllable words, 14 five-syllable words. All analyses investigated the mean number of errors per word. In terms of general performance, the authors did not find any difference between the two groups, but they reported some discrepancies in terms of qualitative errors. Like Whitehouse et al. (2008), they found that the participants with SLI performed worse (in terms of the number of phonemic errors made) than participants with ALI when stimuli were four syllables in length. They suggested, in conclusion, that the claim for a phenotypic overlap between SLI and ALI might have been overstated. It is interesting to note that even if this study can be collocated outside the theoretical frame of Tager-Flusberg and colleagues, the authors used the label ALI for defining children with ASD and language impairment.

Williams et al. (2013): a novel nonword repetition task, in which stimuli were systematically manipulated for length, consonant cluster position and word-likeness, was administrated to 17 HF adolescents with ASD-LI (*M* = 12;38, *SD* = 1.46), 15 adolescents with SLI (*M* = 12;73, *SD* = 2.12) and 19 verbal mental age (VMA) matched TD children (*M* = 6;69, *SD* = 0.33). Children were assessed on eight basic nonword stems containing obstruent + liquid clusters (e.g., /kr, kl, dr, fl/) that were constructed in a 2 (Length: 3 versus 4 syllables) x 2 (cluster position: Initial versus medial) x 2 (morphological suffix: -ist, -ing). A qualitative error analysis was run for each production. Results showed that the SLI group performed significantly worse than the children with ASD-LI, both on general performance and on qualitative error patterns. Essentially the children with SLI were more affected by the position of the consonant cluster in the nonword, were more likely to create novel clusters in incorrect positions of the nonword, and they differed from the participants with ASD-LI in not showing a reliable association between nonword repetition performance and short-term memory capacity evaluated via the short-term memory subtest of the WISC). Crucially, the ASD-LI group did not differ from the VMA-TD group. The authors concluded that, differently from children with SLI, children with ASD-LI perfume as TD children.

Harper-Hill et al. (2013): a group of 24 HF children with ASD aged 11;7 (SD = 28.59 months) and a group of 15 age-matched TD children (*M*= 11;3, *SD* =24.94 months) were assessed through the CNRep task for evaluation of phonological STM. A cluster analysis performed on this task in combination with the Recalling Sentences task (CELF-4) gave rise to two clusters of children within the ASD group: one (ALI) with severe impairment in language abilities that resembled what is found in SLI and another (ALN) with language abilities in line with TD age-peers. However, following Whitehouse et al. (2008), the authors concluded that the similarities between the ALI and the SLI groups might be due to a phenotypical realization of different underlying causes, though no direct motivation for this conclusion were offered in their discussion. Again, this study, like the study by Riches and colleagues, used the ALI label for children with ASD and language impairment without referring to a *continuum* with SLI.

Taylor et al. (2014): The NWR task of the NEPSY- was administrated to a group of 32 HF children with ASD, 61 TD children and 19 children with SLI aged 5- to 12;6- . The children with ASD were further divided into two subgroups, ALI (*n* = 14) and ALN (*n* = 18) on the basis of their performance on the TROG-2 test of receptive morphosyntax. The results showed that the children with ALI and the children with SLI had similar performance on the NWR task. However, no qualitative error analysis was run on the performance of the two groups.

Hill et al. (2015): an experimental nonword repetition test evaluating the proportion of phonemes correctly repeated in nonsense words of increasing length was administrated to 18 children with SLI (*M* = 7;2, SD = 0.8), 22 HF children with ALI (*M* = 6;9, SD = 1) and 20 HF children with ALN (*M* = 6;9, SD = 1.1). The task contained 16 nonwords; 4 stimuli contained one, two, three, and four syllables, respectively. The nonwords were constructed from a limited set of phonemes (11 consonants and 9 vowels) excluding late developing sounds. The nonwords followed an alternating CV – CV structure, and none of the syllables included consonant clusters. Results showed that the both LI groups performed lower than children with ALN. Children with SLI produced significantly more phoneme errors compared to children with ALI, independently of syllable length. These findings reinforced previous studies concluding that the cognitive or linguistic underpinnings of early language difficulties during the early school-age years may be distinct in children with SLI and ALI. However, the authors did not find any correlation between the severity of autism symptomatology and the poor performance of children with ALI, contra Whitehouse et al. (2008). As we said before for Riches and colleagues and Harper-Hill and colleagues, ALI label was used without referring to a *continuum* with SLI.

Tager-Flusberg (2015): four age-matched groups, HF-ALN (*n* = 18; *M* = 10;7 years old), HF-ALI (*n* =20; *M*= 10;1 years old), SLI (*n*=14; *M*=11;2) and TD (*n* = 21; *M* =11;0), were assessed via a nonword repetition task (Dollaghan & Campbell, 1998) with stimuli ranging from 2 to 5 syllables. One-syllable nonwords were CVCs; two-syllable nonwords

CVCVCs; three-syllable nonwords CVCVCVCs; and so on, for a total of 96 phonemes over the entire nonword set. The main findings from this study showed that the children with ALN performed like the TD children and that the children with ALI performed like the children with SLI, with analogous error patterns in both language impaired groups. The authors concluded that the same cognitive mechanisms underlie poor nonword processing in children with SLI and ALI children.

Nadig & Mulligan (2017): 9 HF children with ASD (*M* = 5;7 years old) and 9 MA matched TD children (*M* = 3;0 years old) were assessed on phonological abilities via the Syllable Repetition Task (Shriberg & Lohmeier, 2008) which includes stimuli of 2, 3 or 4 syllables in length, with alternating CV structures. Results showed a significant effect of syllable length, with shorter stimuli being better repeated than longer ones in each of the groups; moreover, the two groups did not differ in repetition accuracy. The authors concluded that children with ASD are not developmentally deviant.

1.2.10.2 General conclusions on NWR tasks

To recapitulate, the key features of the eleven studies reviewed above are summarized in Table 1. From this review we can draw the following conclusions:

1. Putting together all of the studies that evaluated phonological abilities in children with ASD via a NWR task, there was evidence for a subgroup that displays some shortcomings (ASD-LI) and a subgroup that did not show any impairment (ASD-LN). These findings have led some researchers to question whether the subgroup of children with ASD-LI displays the same phenotype as the one characterizing children with SLI:
    a. Tager-Flusberg and colleagues concluded that children with ALI showed the same profile that characterizes children with SLI (Kjelgaard & Tager-Flusberg, 2001; Tager-Flusberg, 2015)
    b. Other researchers have argued tIhat the similarities between children with ASD-LI and children with SLI were only superficial (Taylor et al., 2014); a qualitative error analysis of the two groups highlighted some differences. In general, children with ASD-LI performed better than children with SLI in NWR tasks (Bishop et al., 2004; Botting & Conti-Ramsden, 2003; Harper-Hill et al., 2013; Hill et al., 2015; Riches et al., 2011; Whitehouse et al., 2008; Williams et al., 2013). One study found that children with SLI produced significantly more phoneme errors compared to children with ASD-LI independently of syllable length (Hill et al.,

2015). All the other studies suggested that NWR performance in children with SLI, differently from children with ASD-LI, degraded as syllable length increased: when nonwords were shorter (1 or 2 syllables), children with SLI performed at levels commensurate to non-language-impaired children; when nonwords were longer (≥3 syllables), they performed significantly worse than unaffected children. On the other hand, children with ASD-LI seemed not to be affected by the length of the stimuli, being able to better repeat nonwords with ≥3 syllables. However findings are inconsistent across studies, likely due to methodological differences including the specific features of the particular NWR used (see Table 1). It is important to note that some studies used the ALI label outside the theoretical framework of a possible *continuum* with SLI. We will come back on this topic in section 1.3 and in the discussion of the present work (Part IV).

2. In terms of developmental trajectory, children with ASD-LI did not display long-term phonological impairment, contrary to children with SLI, where deficits are persistent and pervasive (Nadig & Mulligan, 2017; Williams et al., 2013).
3. Some studies, such as (Harper-Hill et al., 2013; Whitehouse et al., 2008, 2007), proposed that NWR deficits in children with ASD-LI were associated with the presence of increased severity of autism symptoms and not with the fact that these children could have the same impairment as children with SLI. However, a more recent study (Hill et al., 2015) reported discordant results, finding that the severity of clinician-observed ASD symptoms was not related with the severity of NWR difficulties.
4. The study by Kjelgaard & Tager-Flusberg, (2001) was the only one that explored the possible relation between general intelligence and performance on NWR, including children with either normal or impaired cognitive abilities. The authors underlined a strong positive correlation between level of intelligence and phonological abilities in children with ASD, indicating that phonological impairment seemed to be linked with shortcomings in the cognitive domain. However, they also noticed that a few children did not follow this pattern, showing normal phonological abilities in presence of intellectual impairment.
5. Williams et al.'s (2013) study seems to stand out from the rest of the studies in terms of the conception of the test, with an attempt at manipulating syllable complexity and including both quantitative and qualitative error analyses, and showing that the children with ASD-LI performed better than the children with SLI.

Limitations of previous studies

I. Within the debate on the possible existence of an SLI profile in ASD, all previous studies used NWR tasks that were highly memory based. We know from the literature that increasing syllable length plays a large role in (decreased) performance in most children with SLI, but it does not seem to have the same effect on children with ASD. Short-term memory could not only influence the performance of the two groups, it could also aggravate (for SLI) or mask (for ASD) the impact of phonological complexity. A more linguistically based task, less dependent on short-term memory, controlled for length of nonwords and specifically built on complexity of syllable structures, ought to be able to satisfactorily assess formal language abilities in both clinical groups.

II. Except for one case (Williams et al., 2013), no study ran a comparison between errors made by ASD-LI and SLI groups. A qualitative error analysis is needed along with general analysis of groups' performance in order to verify the actual phenotypical realization of production in SLI and ASD.

III. A clear picture of the relationship between phonological impairment and severity of autism symptoms did not emerge from the literature: some studies found significant negative correlations between the two measures while others found no correlations. Even more limited is our knowledge about the relationship between phonological abilities and cognitive level in children with ASD (only one study included LF children).

IV. The same labels were used with very different meanings for indicating different populations, contributing to general confusion in the literature.

We will explain in detail in Chapter 4, how the NWR task used in our study addressed these limitations.

TABLE 1. SUMMARY OF THE KEY FEATURES FOR EACH OF THE ELEVEN STUDIES THAT USED A NWR TASK

Study	Nonword Task	ASD group (*N*, age)	Criteria for division between LI and LN subgroups	IQ level	Control group (*N*, age)	Results	Relation with severity of autism symptomatology	Conclusions
Kjelgaard & Tager-Flusberg (2001)	NEPSY	89 children with autistic disorder (DSM-IV); mean age 7;4 years old	Composite score CELF-III and PPVT	FSIQ range 25-141 standard score NVIQ range 43-153 standard score Differential Abilities Scale (DAS)	None	40/89 children could complete the NEPSY NWR task; 34/40 = ≤ -1SD (profile similar to children with SLI)		ALI = SLI
Botting & Conti-Ramsden (2003)	CNRep (Gathercole & Baddeley)	13 children with ASD; mean age 10;10 years old		PIQ > 70 standard score of the WISC-III	29 age-matched children with SLI 25 age-matched children with PLI	SLI < ASD = PLI		ASD with language impairment (ASD-LI) ≠ SLI
Bishop et al (2004)	Baddeley, Gathercole & Watson	80 children with autism and PDD-NOS (ADI-R); mean age 10;4 years old 27 children with speech phrase 53 children without speech phrase		VIQ = 75.9 PIQ = 82.78 WISC-III	59 age-matched TD children	11/80 were not able to completed the task; 27/27 children with speech phrase showed no delay; 42/80 children without speech phrase showed a consistent impairment in the task		ASD without speech phrase ≈ SLI only on a superficial level

Whitehouse, Barry & Bishop (2008)	NEPSY	34 children with autism (DSM-IV) = 18 children with Apoor and 16 children with Aapp	Structural language score derived from the CCC-2	NVIQ ≥ 80 standard score WASI	34 age-matched children with SLI	SLI < Apoor < Aapp on a qualitative error analysis	NWR deficits are associated with more important severity of autism symptoms.	Apoor ≠ SLI
Riches et al. (2011)	CNRep and NMT (Gathercole & Baddeley)	16 adolescents with ALI; mean age 14;8 years old	CELF-III	PIQ ≥ 80 standard score WISC-III	13 age-matched children with SLI 17 age-matched TD children	SLI < ALI < TD on a qualitative error analysis		ALI ≠ SLI
Williams et al. (2013)	Experimental NWR task	17 adolescents with ASD-LI mean age 12;3 years old	CELF-4 (cut-off 78)	PRI > 80 WISC-IV	15 age-matched children with SLI 20 CA matched TD children 19 VMA matched TD children	SLI < ASD-LI = VMA-TD < CA-TD		ASD-LI ≠ SLI ASD-LI are developmentally delayed
Harper-Hill, Copland & Arnott (2013)	CNRep (Gathercole & Baddeley)	24 children with ASD; mean age 11;7 years old = 6 children ALI and 18 children ALN	CELF-4	RPM score whithin the norm	15 age-matched TD children	ALI < ALN = TD	NWR deficits are associated with more important severity of autism symptoms.	ALI ≠ SLI no direct explanation was given for this conclusion
Taylor et al. (2014)	NEPSY	32 HF children with ASD aged 5- to 12;6-years old; ALI (n = 14), ALN (n = 18)	TROG-2 receptive morphosyntax	WASI matrix reasoning (cut-off - 1 SD)	61 TD children and 19 children with SLI	ALI = SLI		

Hill et al. (2015)	NTR	42 children with ASD (DSM-IV-TR) = 22 children with ALI (*M* = 6;9) and 20 children with ALN (*M* = 6;9)		PRI > 80 WIPPSI or WISC-IV	18 age-matched children with SLI	SLI < ALI < ALN on a qualitative error analysis	NWR deficits are **not** associated with more important severity of autism symptoms.	ALI ≠ SLI
Tager-Flusberg (2015)	Experimental NWR task	38 children with ASD (DSM-IV) = 18 children with ALN (M=10;7) and 20 children with ALI (M = 10;1)		FSIQ of the Differential Abilities Scale within the norm	14 age-matched children with SLI 21 age-matched TD children	SLI = ALI < ALN = TD		ALI = SLI
Nadig & Mulligan (2017)	SRT (Shriberg et al., 2009)	9 children with ASD		NV skills of the Mullen Scales of Learming Abilities within the norm	9 mental age-matched TD children (M = 3; 0 years old)	ASD = MA-TD		ASD are developmentally delayed

1.2.11 Sentence Repetition (SR) Tasks

Sentence Repetition (SR) is gaining increasing attention as a source of information about children's sentence-level abilities in clinical assessment. Specifically, one of the earliest studies using this test as a possible endophenotype of language impairment (Conti-Ramsden et al., 2001) showed that SR is very sensitive in identifying SLI. SR is not a mere verbatim echoing of the stimulus: if adequately modelled in order to minimize the effect of short term memory and to include an adequate degree of syntactic complexity, it can evaluate the effects of different types of long-term linguistic knowledge on immediate recall (Lombardi & Potter, 1992; Potter & Lombardi, 1990; Potter & Lombardi, 1998). To perform this type of task, participants have to comprehend the sentence in terms of their abstract grammatical system representation and then process the linguistic information using their own grammatical and memory systems (Baddeley, 2000).

While SR should reveal impairment in the domain of morphosyntax, it should be less sensitive to pragmatic deficits (Polišenská et al., 2015). The design of SR incorporates formal aspects of morphosyntactic processing but is less constrained by pragmatic features (e.g. giving an answer related to a context or using inference to complete a sentence, a.o.). In SR children are instructed to repeat sentences exactly as heard without any given pragmatic context. In other words, formal language abilities and the ability to use language in context are not densely intertwined in SR, in contrast to what is typically the case in other kinds of tasks (act-out, sentence-picture matching, sentence completion, a.o.). Therefore following Botting & Conti Ramsden, a.o., who proposed to use this task in ASD as a good endophenotype for language impairment, we hypothesised that SR should also narrow the possible influence of pragmatic impairment in this pathology, to investigate children's morphosyntactic abilities and to reveal possible similarities between ASD and SLI (Silleresi et al., in press).

However, not all SR tasks are designed in the same way. Depending on its structure, SR can evaluate essentially short-term memory or more specifically morphosyntactic constructions built on computational complexity. Computational complexity, in a generative linguistic theoretical framework, can be measured in terms of the nature and number of operations needed for the derivation of a syntactic construction (Jakubovicz & Tuller, 2008). This account suggests that children with language impairment, notably children with SLI, show a deficit in the computational system, which leads to the inconsistent use of certain grammatical operations. It has been demonstrated that structure-dependent relationships that can be observed in tense marking, case marking and long-distance dependencies which necessitate movement may constitute a source of difficulties for these children (Franck et al., 2004; Hamann, 2006; Jakubovicz & Tuller, 2008, a.o.).

In the next section we will report on the very few studies (seven) that have employed SR for evaluating morphosyntactic abilities in children with ASD. We will see that the vast majority (6/7) of the studies that compared children with ASD to TD age-peers and/or children with ASD-LI to children with SLI employed tasks that are highly memory based. Five of these used the Recalling Sentences task of the CELF. This task evaluates the ability to repeat sentences of increasing length and complexity *verbatim*. It is composed of 32 stimuli including a variety of different morphosyntactic constructions (active declaratives, wh-questions, passives), but sentence length plays a large role in this task. The length of the sentences dramatically increases from 6/7 words (7/8 syllables) in the first sentences, to 14 words (15 syllables) around halfway through the task, and up to 18 words (23 syllables) in the last sentences. One study employed the Sentence Repetition subtest of the NEPSY-II, which it is composed of 17 sentences. This task, which is part of the "Memory and Learning" index of the NEPSY battery, and, crucially, not part of the "Language" index, was constructed to assess verbal memory span and short-term memory. The manual of the NEPSY states that: "the subcomponents assessed include working memory and immediate and delayed recall of simple stimuli, unprompted descriptive memory, cued recall, and recognition, repetition, and recall amid interference. A low score suggests poor verbal short-term or immediate memory for meaningful sentences". Finally, only one study used a SR task built specifically to evaluate morphosyntactic complexity. However we will see that this task also shows limitations.

Results will highlight that the ASD-LI/SLI parallelism can be inferred only at a very general level. In the next section we will outline the main findings of the studies that evaluated morphosyntactic abilities via SR tasks. At the end of the section we will pinpoint the strengths and the limitations of these studies and argue for the use of a specific SR task, the LITMUS-SR, for a purer evaluation of morphosyntactic abilities in children with ASD.

1.2.11.1 Literature review on studies that employed SR tasks

To our knowledge, only seven studies have employed SR as a possible marker of syntactic deficits in ASD.

Botting & Conti-Ramsden (2003): they used several subsets of the CELF-R, including the Recalling Sentences subtest, to evaluate morphosyntactic abilities in 13 HF children with ASD, 29 children with SLI and 25 children with Pragmatic Language Impairment (mean age 10;10). The SR task turned out to be the most accurate marker for differentiating the three groups of children with communication disorders and highlighted similar performance between the children with ASD-LI and the children with SLI, who both performed in the impaired range.

Loucas et al. (2008): the Recalling Sentences subtest of the CELF-R was employed, among other standardized tests, to assess syntactic abilities in 41 HF children with ALI (*M* = 11;4 years old), 31 HF children with ANL (*M* = 11;7 years old.), and 25 children with SLI (*M* = 12.7 years old). Results showed that the performance of the two impaired groups did not differ, while the ANL group performed significantly better. The authors suggested that the linguistic performance of autistic children was not related to their level of autism severity. No qualitative error analysis was run on the results.

Riches et al. (2010): this was the only study which included an experimental task of sentence repetition, in conjunction with the Recalling Sentences subtest of the CELF-III, in three groups of HF adolescents (*M* = 14;4-15;3 years old) with ALI (*n* = 16), SLI (*n* = 14) and TD (*n* = 17). The experimental task included a set of 24 sentences containing relative clauses modelled on four different configurations of increasing complexity. Consistent with previous research, sentence repetition demonstrated sensitivity as a marker of language impairment, yielding high error rates in both clinical groups and low error rates in the TD participants. However, the participants with SLI were significantly more affected by syntactic complexity, and more prone to make wholesale changes to the model syntactic structure, than the children with ALI. Wholesale changes included the transformation of object relatives into subject relatives and viceversa, and the production of passive object relatives instead of object relative sentences. Children with ASD-LI did not show different error types than children with SLI, but only lower error rates.

Harper-Hill et al. (2013): they used the same task as the one used in Botting & Conti-Ramsden (2003), Loucas et al. (2008) and Riches et al. (2010), i.e. the CELF-4 Recalling Sentences subtest, in a group of 24 English-speaking children aged 9- to 16- years with a diagnosis of HF-ASD. Results showed the existence of two clusters of children, the first one (ALI) with severe impairment in language abilities that resembled what is found in SLI and a second cluster of children (ALN) with language abilities in line with TD age-peers. The first cluster was strongly associated with greater severity of autistic symptomatology. However, following Whitehouse et al. (2008), the authors concluded that the similarities between the ALI and the SLI groups might have been due to a phenotypical realisation of different underlying causes, though no direct motivations for this conclusion were given.

Taylor et al. (2014): The SR task of the NEPSY-II was administrated to a group of 32 HF children with ASD, 61 TD children and 19 children with SLI aged 5- to 12;6-.Children with ASD were further divided into ALI (*n* = 14) and ALN (*n* = 18) on the basis of their performance on the TROG-2 test of receptive morphosyntax. The results showed that the children with SLI performed significantly worse than the children with ALI on the SR task. However, no qualitative error analysis was run on the performance of the two groups.

Schaeffer (2016): through use of the Recalling Sentences subtest of the CELF-4, she compared the performance of 27 children with HFA and no formal language impairment (*M* = 10;0 years old) with 27 age-matched children with SLI and 27 age-matched TD children. Results showed that the ASD group did not differ from the TD controls, while the SLI group performed significantly lower.

Brynskov et al. (2017): two groups of 12 HF children with ASD with delay (*M* = 6;1 years old) and 9 without delay (*M* = 6;4 years old) in first word production, were presented with the Recalling Sentences subtest of the CELF Preschool-2. Their performance was compared to that of a group of 21 slightly younger TD children (*M* = 5;5 years old). Results showed that while the ASD+Delayed subgroup had impaired performance on sentence repetition, the ASD-No Delay subgroup displayed intact performance, comparable to the TD group. This study was the only one that clearly stated that since the children with the strongest sentence repetition performance also had the most advanced comprehension skills and did not show any evidence of immediate echolalia in their spontaneous speech, their performance on sentence repetition did not appear to be fully automatic or echolalic.

1.2.11.2 General conclusions on SR tasks

To recapitulate, we summarized the key features of the seven studies reviewed above in Table 2. From this review we can set the main features of the current debate as follows:

1. All studies identified SR as a very sensitive tool for detecting formal language capabilities in children with ASD.
2. There is evidence that among children with ASD there is a subgroup that displays morphosyntactic impairment (ASD-LI) and a subgroup that does not show any such impairment (ASD-LN).
3. Among the five studies that directly compared children with ASD-LI and children

with SLI on SR tasks, three found similar performance between the two impaired groups (Botting & Conti-Ramsden, 2001; Loucas et al., 2008; Harper-Hill et al., 2013), while the other two did find more impaired performance in children with SLI (Riches et al., 2010; Taylor et al., 2014).

4. Harper-Hill et al., (2013) was the only study that proposed that low SR performance in children with ASD-LI might be associated with greater severity of autism symptoms.
5. All of these studies investigated morphosyntactic abilities in high-functioning individuals.
6. One study evoked the possibility that echolalia might not be a facilitating factor in the sentence repetition task for children with ASD.

Limitations of previous studies

I. None of the seven studies utilized a linguistically based sentence repetition task which targeted computational complexity and included, at the same time, a variety of different structures. A detailed explanation of the reason why both of these features should be included in the task will be given in Chapter 5. For now we can say that, first of all, the main limitation of the Recalling Sentences subtest (CELF) and of the SR task of the NEPSY-II is the fact that both these tasks are highly memory based. It is very difficult to separate performance on complex structures from the relationship with short-term memory in this kind of evaluation. Secondly, while the SR task in Riches et al. (2010) met the criterion of computational complexity, it focused only on relative clauses, which prevents investigation of a variety of complex constructions and excludes simple sentences altogether. Having other constructions than relatives in the SR task could be useful for detecting potential difficulties and errors in groups of children with language impairment and therefore be useful in comparing children with ASD-LI and SLI. Moreover a more heterogeneous task would be easily transferable to clinical practice since it should be less stressful for children, who should at least be expected to succeed in repeating less complex sentences. Finally a task including both simple and complex sentences should be more likely to be useable with a wide variety of children with ASD along the spectrum. In Chapter 4 we will explain how the SR task used in our study addressed these limitations.

II. As is the case for NWR, as seen above, qualitative error analysis is needed along with general analysis of group performance in order to verify the actual phenotypical characteristics of production in SLI and ASD-LI.
III. There is a significant lack of results on the relationship between morphosyntactic impairment and severity of autism symptoms.
IV. No study included LF children with ASD.

TABLE 2. SUMMARY OF THE KEY FEATURES FOR EACH OF THE SEVEN STUDIES THAT USED A SR TASK

Study	Sentence Repetition task	ASD group (*N*, age)	Criteria for division between LI and LN subgroups	IQ level	Control group (*N*, age)	Results	Relation with severity of autism symptomatology	Conclusions
Botting & Conti-Ramsden (2003)	Recalling Sentences of the CELF-R	13 children with ASD; mean age 10;10 years old		PIQ > 70 standard score of the WISC	29 age-matched children with SLI 25 age-matched children with PLI			ASD-LI = SLI
Loucas et al. (2008)	Recalling Sentences of the CELF-R	72 children with ASD = 41 children with ALI (mean age 11;4 years old) and 31 children with ANL (mean age 11;7 years old)	CELF-III Receptive Language, Expressive Language or Total Language score of ≤ 77	PIQ > 80 standard score WISC-III	25 slightly older children with SLI (mean age 12;7)	SLI = ALI < ANL		SLI = ALI
Riches et al. (2010)	Experimental task of SR (Relative Clauses) + Recalling Sentences of the CELF-III	16 adoelescents with ALI; mean age 14;8 years old	CELF-III Receptive Language, Expressive Language or Total Language score of ≤ 77	PIQ > 80 standard score of the WISC-III	14 age-matched children with SLI 17 age-matched children with TD	SLI < ALI < TD on a qualitative error analysis		ALI ≠ SLI

Harper-Hill et al. (2013)	Recalling Sentences of the CELF-4	24 children with ASD; mean age 11;7 years old = 6 children ALI and 18 children ALN	CELF-4	RPM score whithin the norm	15 age-matched TD children	ALI < ALN = TD	SR deficits are associated with more important severity of autism symptoms.	ALI ≠ SLI no direct explanation was given for this conclusion
Taylor et al. (2014)	SR NEPSY-II	32 HF children with ASD aged 5- to 12;6- years old; ALI (n = 14), ALN (n = 18)	TROG-2 receptive morphosyntax	WASI matrix reasoning (cut-off - 1 SD)	61 TD children and 19 children with SLI	ALI < SLI		
Schaeffer (2016)	Recalling Sentences of the CELF-4-NL	27 children with HFA (and no fromal language impairment); mean age 10;0 years old	CELF-4 NL	NVIQ > 80 standard score on Raven Progressive Matrices	27 age-matched children with SLI 27 age-matched children with TD 16 TD adults; mean age 34;2	SLI < HFA = TD children		HFA ≠ SLI
Brynskov et al (2016)	Recalling Sentences of CELF-R	21 children with ASD; mean age 6;1 years old = 9 children ASD-No Delay subgroup and 12 children ASD + Delay subgroup	children produced single words by age two and sentences by age three	Raven Progressive Matrices scaled scores > 80	21 age-matches TD children	ASD + Delay < ASD -No Delay = TD		ASD + Delay < ASD -No Delay = TD

1.3 ASD-LI and ASD-LN: labels for whom?

It is a well-known fact that there are at least two profiles of formal language abilities in autism, one with normal formal language abilities (LN) and one with impaired formal language abilities (LI). This section is meant to illustrate the labelling issue that we raised early (notably the fact that the labels for these two subgroups have been quite inconsistent over time and dependent on the linguistic measures used for the identification of language impairment). Table 3, reports on the criteria that have been adopted for dividing children with ASD into the LI and LN subgroups, the names that have been assigned to these linguistic profiles, the cognitive level of the targeted population, and finally, the conclusions that have been reached on the similarities between the LI profile and SLI and between the LN and TD children. From Table 3, we can observe:

The heterogeneity of the labels

Researchers have used many labels for defining profiles of structural language abilities in children with ASD. We have identified at least four different pairs of labels used to distinguish LI and LN profiles (ASD-LI/ASD-LN; ALI/ALN; Apoor/Aapp; ASD + Delay/ASD No delay).

The heterogeneity and inconsistency of tools and criteria

The tools and the criteria through which children with ASD were divided into profiles of structural language abilities varied from study to study (performance on receptive vocabulary; performance on receptive and expressive vocabulary; composite score of vocabulary, morphosyntax and phonology; composite score drawn from standardized tests, like the CELF, targeting vocabulary, morphosyntax, phonology, semantics; subtest of morphosyntactic abilities and working memory, like the RS of the CELF; subtest of phonological abilities; developmental factors such as age of first word and/or first sentence, etc.).

Different cut-offs for LI have been used for the same tests through the studies (e.g. Williams et al., 2013 vs. Hill et al., 2015), yielding further confusion.

Concerning the use of standardized tests and experimental tasks for dividing children into LI and LN profiles, it might be the case that some children were placed in the LI subgroup due to them having misunderstood the pragmatics of the testing situation rather than having real impairment in structural language.

Population heterogeneity

The same labels (ALI and ALN) have been used to identify sometimes only children with HF autism and sometimes groups including both children with LF and HF autism.

Heterogeneity of conclusion

Firm conclusions concerning the similarities between the LI and SLI profiles and the LN and TD profiles cannot be drawn, since the same labels (ALI and ALN) were used both by studies that found strong similarities and by studies which highlighted strong differences.

We can, therefore, ask ourselves what the LN and LI labels really indicate and crucially which kind of conclusions we can draw from all of the studies presented in section 1.2. Since studies have divided children into LI and LN groups with different tools and criteria and have included different populations, how trustworthy can the comparison of results be? It is evident that all of the above-mentioned features need more controlled selection and evaluation.

Regarding the heterogeneity of the labels, after the literature review reported in section 1.2 and in Table 3 we can now come back to our choice of using ASD-LI and ASD-LN in reference to the two subgroups of formal language abilities in children with autism. With the aim of maintaining neutrality we chose not to use the ALN and ALI labels. The reasons were twofold: first, these labels were introduced by Tager-Flusberg and colleagues with the precise intent of referring to ASD and SLI as a *continnum* of impairment. Second, the fact that the meaning of ALI and ALN fluctuates from one study to another (sometimes indicating phenotypical similarities between children with ALI and SLI and sometimes not) is a source of confusion.

Labels referring to developmental language milestones such as age of first word and/or sentence were based on the results reported in Kover et al., (2016), who run a large longitudinal study supporting the idea that age of language onset is only partially predictive of later linguistic development in children with ASD. Similarly we did not agree with the choice of dividing children with ASD between Apoor and Aapp on basis of the scores obtained on a parental questionnaire, since there was no direct evaluation of children's language abilities. Our choice of using ASD-LI and ASD-LN was also motivated by the fact that these labels were assigned in some studies, notably Botting & Conti-Ramsden (2003) and Harper-Hill et al. (2013), as *a posteriori* descriptions of structural language abilities of children with ASD, without any specific reference to SLI or Typical Development. This choice avoids all the *a priori* limitations listed above.

TABLE 3. AN OVERVIEW OF THE TERMINOLOGY, TOOLS, AND POPULATIONS PREVIOUSLY USED TO DEFINE CHILDREN WITH ASD WITH AND WITHOUT FORMAL LANGUAGE IMPAIRMENT

Labels	Studies	Criteria for LI profile	Cognitive level of the population	Conclusions
ALI and ALN	Kjelgaard & Tager-Flusberg 2001	Below the norm of the PPVT and/or the CELF-III	LF and HF	ALI = SLI
	Roberts et al. 2004	Below the norm of the PPVT	LF and HF	ALI ≠ SLI
	Tager Flusberg 2006	Below the norm on EVT (expressive vocabulary) and PPVT (receptive vocabulary)	HF	ALI = SLI
	Loucas et al. 2008	CLEF-III total score < 77	HF	ALI < SLI on standardized tests; ALI = SLI on SR
	Riches et al. 2010	CELF-III total score < 77	HF	ALI > SLI
	Riches et al. 2011	CELF-III total score < 77	HF	ALI > SLI
	Taylor et al. 2014	TROG-2 < -1 SD	HF	ALI = SLI on NWR; SLI < ALI on SR
	Hill et al. 2015	CELF-4 total score < 85	HF	ALI = SLI
	Tager Flusberg 2015	CELF-III total score < - 1 SD	HF	ALI = SLI
	Perovic et al. 2013; Modyanova et al. 2017	Below the 10th percentile on TROG2 + below the 10th percentile in at least KBIT/ PPVT	LF and HF	ALI < SLI and ALN < TD
	Ellis Weismer et al. 2017	CELF 4 at least -1,25 SD below the mean on one or more summary scores from the CELF (core language, expressive language, receptive language)	HF	No direct comparison between ALI and SLI

Labels	Studies	Criteria for LI profile	Cognitive level of the population	Conclusions
Apoor and Aapp	Whitehouse et la. 2008	Structural language score derived by the CCC2	HF	Apoor > SLI
Language onset	Bishop et al. 2004	Language onset: without speech phrases at 36 months (subscore ADI-R)	HF	ASD with Speech delay are similar to SLI only on a superficial level
ASD with delay/ ASD + No delay	Brynskov etal. 2016	Language onset: if they did not produce single words at age 2 and sentences at age 3	HF	ASD + No delay = TD
ASD-LI and ASD-LN	Botting & Conti Ramsden 2003	Recalling Sentences of CELF-III and CNRep	HF	ASD and LI = SLI on RS; ASD and LI > SLI on CNRep
	Harper-Hill et al. 2013	Recalling Sentences of CELF-4 and CNRep	HF	ASD and LI ≠ SLI
	Williams et al. 2013	CELF-4 total score < 78	HF	ASD-LI > SLI
	Tuller et al. 2017	< 1.65 SD on a composite score of vocabulary (ELOLA), morphosyntax and phonology (BILO)	LF and HF	ASD-LI = SLI and ASD-LN < TD on a general performance
	Durrleman et al. 2017	< 2 SD on a composite score of vocabulary (ELOLA) and morphosyntax (NEEL)	LF and HF	ASD-LI = SLI and ASD-LN < TD on a general performance

1.4 Conclusions and research direction for the present work

At the end of the present chapter we can draw the following conclusions. Studies focusing on formal language in ASD have found two profiles of abilities (LI and LN) both for phonology and morphosyntax. Concerning morphosyntax, the phenotypical realisation of language abilities in ASD-LI has been found to be similar to SLI, while language capacities in the ASD-LN profile have compared to TD acquisition. One of the main issues raised in the literature was whether the similarities found between ASD-LI and SLI are genuine or whether they may result from different underlying sources of impairment. Some studies have, in fact, identified qualitative differences in error patterns between these two populations, suggesting that language errors or unexpected productions in children with ASD-LI might be due to pragmatic shortcomings (the children fail to understand the conversational situation) rather than to a linguistic breakdown similar to the one found in SLI. Moreover, a few studies found qualitative differences between children with ASD-LN profile, in comparison to TD children, raising the question as to what extent language abilities in children with ASD-LN are really spared.

In turn, although phonological abilities in children with ASD have generally been indicated as spared, the very few studies that compared ASD-LI to children with SLI, concentrating on qualitative error analyses, have highlighted some striking similarities between these two groups of children. It has been questioned, then, whether supposedly spared performance of children with ASD on tasks of articulation and repetition of real words might be a result not only of phonological knowledge, but of several other factors, such as pre-existing lexical knowledge, which in relation to word familiarity or frequency effect may create biased performance in these children. Some researchers supposed that children with ASD having good vocabulary knowledge could rely on their lexical abilities rather than on phonological capacities, in performing on these kinds of tasks.

Disentangling language impairment in ASD is not an easy task, since for establishing the nature of language difficulties of children with ASD it is not always possible to tease apart language difficulties due to formal language impairment and those related to problems with pragmatics or word familiarity effects. In particular, measures drawn from standardized tests rarely enable detailed analysis of specific morphosyntactic and phonological structures. Moreover, these tests generally target multiple aspects of several language domains such as morphology, lexicon, syntax, phonology and articulation, at the same time, preventing distinct evaluation of these abilities.

In parallel some researchers started to use repetition tasks, SR and NWR, for evaluating formal abilities in autism. The hypothesis underlying this choice was twofold: (1) since these tasks have been proved to be very sensitive to language impairment in clinical assessment of SLI, they may also be able to detect structural language deficit among children with ASD. (2) The structure of both tasks should be ideal for an evaluation of formal language abilities in children with ASD, since they both highlight underlying language impairment without concurring factors. Pragmatic impairment typical of children with ASD should have limited impact on repetition accuracy in SR tasks. The design of SR incorporates formal aspects of morphosyntactic processing but is less constrained by pragmatic features. Similarly, the design of a NWR task should ensure a more controlled evaluation of phonological abilities, since individuals cannot rely on their pre-existing lexical knowledge, when repeating nonwords conceived to be unrelated to existing words in the child's language. The few studies which have used repetition tasks in children with autism found that these tasks distinguished the same two subgroups of language abilities (ASD-LI and ASD-LN) both on SR and NWR. Nonetheless, when compared with children with SLI, children with ASD-LI displayed sometimes identical patterns and sometimes different patterns of morphosyntactic and phonological abilities, depending on the study. When we looked at the types of SR and NWR tasks used in these studies we noticed some methodological limitations. Essentially, all tasks were highly memory based and not specifically constructed for evaluating linguistic abilities.

Moreover, the relationship between language abilities and extra-linguistic factors such as, cognitive abilities and severity of autism symptoms, has not yet given rise to a complete picture in the literature. Since very few studies have investigated the link between these two extra-linguistic factors and formal aspects of language, it remains to be seen if generalized learning disability and/or severity of autism symptoms might contribute to poor performance of children on formal aspects of language or whether these results could be caused from a specific linguistic breakdown, as is the case for children with SLI (see Jakubowicz, 2005 for an overview).

Moving from these conclusions we can identify the following directions of research for the present work:

(1) Our literature review highlighted several methodological problems in evaluating structural language abilities in children with ASD both for morphosyntax and

phonology. Following the idea that SR and NWR task should be the most adapted tools for evaluating structural language abilities in children with autism, we propose to further specify the evaluation of structural language abilities using linguistically based tasks for comparing performance of children with ASD to performance of children with SLI and TD children. In the present study we sought to evaluate morphosyntactic and phonological abilities in French-speaking children with ASD, through use of the French version of the Language Impairment Testing in a Multilingual Setting, LITMUS Sentence Repetition task (LITMUS-SR) and Nonword repetition task (LITMUS-NWR), developed for detecting SLI in monolingual and bilingual children, within COST Action IS0804.[5] These tests were specifically created to be linguistically based and not memory based.

(2) No clear conclusions have been drawn on the similarities/differences between ASD-LI and SLI and between ASD-LN and TD. We propose to further analyze this question, integrating a quantitative and qualitative analysis of results. Moreover, we propose to search for possible differences in developmental trajectory.

(3) Very few studies have specifically investigated the interaction of extra-linguistic factors and the findings have been quite mixed

We propose to investigate at what degree extra-linguistic factors, notably low cognitive abilities and severity of autism symptoms, have an effect on structural language abilities

[5] COST Action IS0804 (2009-2013) developed a series of tasks for Language impairment testing in multilingual settings (LITMUS); see www.bi-sli.org.

Chapter II
Cognitive profiles in Autism Spectrum Disorder

2.1 Introduction

To date, few studies have taken up the question of the nature of linguistic/cognitive profiles in children with ASD by explicitly exploring the interaction between language (dis)ability and intellectual (dis)ability and therefore their logically possible combinations. Typically, researchers have found evidence for three profiles of structural language (phonology and/or morphosyntax) and intellectual abilities: two "homogeneous" profiles, ASD-LN with HF abilities and ASD-LI with LF abilities and one "discrepant" profile, ASD-LI with HF abilities, similar to the profile that defines children with SLI. However, some studies have evoked the existence of another "discrepant" profile, i.e. ASD-LN with LF abilities (Joseph et al. 2002; Kjelgaard & Tager-Flusberg, 2001; Tuller et al. 2017). The possible existence of four linguistic/cognitive ability profiles raises several theoretical considerations. First, it undermines the old belief that HF abilities necessarily lead to normal language abilities, and that LF abilities entail *ipso facto* impaired formal language skills. Moreover, from the vantage point of a language module in the human mind/brain, which thus can be selectively spared (Fodor, 1985) (even though this module interfaces nonetheless with other modules and central systems) these results receive a natural interpretation. Adapting the words of Smith & Tsimpli (1995), the existence of these varied profiles provides a classical example of *double dissociation*: language can be impaired in some children with ASD of otherwise normal intelligence, and—more surprisingly—some children with ASD and impaired intelligence may nonetheless have normal, or even enhanced, linguistic ability. The existence of a profiles like the one found in the ASD-LN with LF cognitive abilities indicates that children with ASD can indeed display spared language abilities in the presence of impaired nonverbal intelligence, a profile reminiscent of that found in Williams Syndrome (Mervis & Velleman 2011) and also in the well-known case study of the Savant *Christopher* (Smith & Tsimpli, 1995). We will come back to these topics in Part III (bilingual children with ASD) and in the general discussion of the present work.

As we saw in Chapter 1, the relative infrequency of studies investigating formal language abilities in ASD that have included children with low cognitive profiles in their population samples (16 studies in total, with sometimes very few such children) means that there is a lack of knowledge about the capabilities of LF children, and especially the ASD-LN with LF ability profile. Four of these studies found evidence for all four structural language/cognitive profiles in their population samples (Kjelgaard & Tager-Flusberg, 2001; Joseph et al. 2002; Tuller et al. 2017). Tuller and colleagues highlighted that a real understanding of linguistic/cognitive profiles in autism can only be achieved through investigation of the entire spectrum, which is not restricted to high-functioning autism.

Similarly to our conclusion that repetition tasks should be the most suitable tools for evaluating structural language abilities in children with ASD, among the ones proposed in the previous chapter, the main aim of this chapter is to see whether it is possible to identify the most appropriate tool(s) for evaluating cognitive abilities in relation to linguistic abilities in ASD population. What do we mean by *most appropriate*? For the purpose of the present study, and more generally for an evaluation of cognitive profiles in relation to linguistic abilities in children with ASD, a well-suited tool should meet the following criteria:

I. A tool that is easily available and that can be easily used both in research and clinical practice.

II. A tool that identifies cognitive abilities of individuals on the spectrum, without any confounding effects of other variables, e.g. fine motor or speech skills (much like what repetition tasks did for structural language abilities).

III. A tool that is, at worst, only remotely related to language abilities. Since the purpose of the present work is to detect profiles of language and cognitive abilities in children with ASD, our desire is to use cognitive scores which are as nonverbal as possible, just as we have sought to use language measures capable of measuring as faithfully and specifically as possible structural language abilities.

Do such tools exist? In the next section we will report on the main characteristics of intelligence in autism and pinpoint the tools that have been argued to be the most appropriate to evaluate cognitive abilities in children with ASD. We will then verify whether these tools could be used for the purposes of the present study, checking whether they satisfy all four criteria listed above.

2.2 What do we know about intelligence in autism?

The observation that intelligence in autism is atypical dates back to its earliest descriptions (Asperger, 1944; Kanner, 1973) with empirical studies reporting uneven autistic performance both across and within commonly used intelligence tests (Bartak et al., 1975). It is now well established that while for typical population samples the use of different psychometric tools and the evaluation of different cognitive domains provide the same results on intellectual abilities (see Mackintosh & Mackintosh, 2011 for a review), this does not hold for individuals with ASD. The difficulty with employing intelligence measures in autism is that IQ profiles are not flat in this pathology. Individuals with autism typically show peaks (strengths) and valleys (weaknesses) of performance on specific domains of cognitive abilities on the various subtests of most, if not all, intelligence test batteries (Mottron, 2004). Peaks and valleys can be identified in either absolute or relative terms. A relative peak/valley is an area in which an individual excels/fails compared with other areas of cognitive abilities in which (s)he performs worse/better. An absolute peak/valley is an area in which an entire population or group of individuals excels/fails, compared with another group of individuals who perform worse/better (Stevenson & Gernsbacher, 2013).

There is evidence that children with autism tend to display a relative and sometimes absolute (when compared to TD age-peers) weakness on tasks evaluating attention, motor abilities, processing speed and verbal abilities, especially those which entail measures of social and practical understanding and/or pragmatic-communicative skills based on culture-specific demands (Minshew et al., 2002). On the other hand, nonverbal (NV) reasoning has been demonstrated to be a relative strength of autistic individuals. Children (and more generally, individuals) with ASD tend to display peaks of performance on nonsocial materials, notably tasks based on abstract perception and visuo-spatial abilities (Mottron et al., 2006). NV tests, in fact, typically minimize the need for task instructions, culture- or experience-specific abilities, and other specific abilities which may be important for performing the task (e.g., fine motor or speech skills). Moreover, since NV tests generally target abstract and fluid reasoning, children with ASD can leverage on their enhanced abstract spatial processing, perceptual and visuospatial skills to perform these tasks (Courchesne et al., 2016).

Two paradigmatic examples of these peaks and valleys of abilities are represented by (1) the difference in performance of the ASD population on the two main instruments used to evaluate human intelligence, notably Raven’s Progressive Matrices (RPM) in research and Wechsler Intelligence Scale for Children (WISC) in clinical practice (Stevenson &

Gernsbacher, 2013),[6] and (2) by the difference in performance that children with ASD typically display on the various indices that compose the internal structure of the Wechsler scales. Before reviewing both of these examples of discrepant performance, we present here the main characteristics of these two psychometric tools:

- **Raven's Progressive Matrices (RPM)** is a one-format 36-item test divided into three sets of 12 items, which increase in difficulty and complexity within and across sets. Each item is composed of a pattern with a piece missing, leaving an empty hole or space in the board. There are six possible pieces underneath, among which the one that best completes the matrix must be chosen to fill the empty space. RPM has been identified as the most complex and general single test of intelligence, a metric of reasoning and problem solving, and it is believed to be a ''paradigmatic'' measure of fluid intelligence. The RPM task has been empirically demonstrated to assay the ability to infer rules, manage a hierarchy of goals, and form high-level abstractions (Snow et al., 1984).
- **Wechsler Scales of Intelligence**. Since it will be the version used in the present study, we present here the fourth edition of the Wechsler scale (WISC-IV). This version is composed of ten core subtests grouped into four indices, evaluating different domains of general intelligence, from which a Full Scale IQ (FSIQ) is derived. The four indices are the PRI (Perceptual Reasoning Index), the VCI (Verbal Comprehension Index), the WMI (Working Memory Index), and the PSI (Processing Speed Index). Here we describe the four indices and the ten subtests included in the WISC-IV:
- The **VCI (Verbal Comprehension Index)** assesses verbal reasoning and comprehension via three core subtests:
  - **Similarities** measures logical thinking, verbal concept formation and verbal abstract reasoning. Two similar objects or concepts are presented, and the participant is asked to tell how they are alike or different.

[6] The Wechsler intelligence scales were developed by Dr. David Wechsler in 1939 to measure intellectual functioning of children and adults, and they have been subsequently revised over the years. The Wechsler Adult Intelligence Scale (WAIS) is intended for use with adults. The Wechsler Intelligence Scale for Children (WISC) is designed for children aged 6 to 16, while the Wechsler Preschool and Primary Scale of Intelligence (WPPSI) is designed for children aged 4 to 6 years old.

- **Vocabulary** measures the individual’s vocabulary knowledge, word usage and capacity to define words.
- **Comprehension** measures common-sense social knowledge, practical judgment in social situations, and level of social maturation, along with the extent of development of moral conscience. Children are asked to explain situations, actions, or activities that they'd be expected to be familiar with.

- The **PRI (Perceptual Reasoning Index)** assesses NV abilities, fluid reasoning and visuospatial skills via three core subtests:
  - **Block Design** measures an individual’s ability to analyse and synthesize an abstract design and reproduce that design from coloured plastic blocks. The child must arrange the blocks to match the design formed by the examiner or shown on cards. Spatial visualization and analysis, simultaneous processing, visual-motor coordination, dexterity, and nonverbal concept formation are involved.
  - **Picture Concepts** measures categorical and abstract reasoning. Individuals are asked to look at two (or three) rows of pictured objects and indicate the single picture from each row that shares a characteristic in common with the single picture(s) from the other row(s), explaining (verbally) why the objects chosen go together.
  - **Matrix Reasoning** measures visual processing and abstract, spatial perception. Similarly to RPM, children are shown coloured matrices, or visual patterns, with something missing. The child is asked to select the missing piece from a range of options. Matrix reasoning has been recognized as relatively culturally-fair and language free. Moreover, it does not require any direct hand-manipulation.

- The **WMI (Working Memory Index)** measures the ability to register, maintain, and manipulate visual and auditory information in conscious awareness, which requires attention and concentration, as well as visual and auditory discrimination. It is composed of two subtests:
  - **Digit Span** measures short-term auditory memory and attention. The digits have no logical relationship to each other and are presented in random order by the examiner. The individual must then recite the digits correctly by recalling them forward (FDS) and backwards (BDS).

- **Letter-Number Sequencing** measures attention span, short-term auditory recall, processing speed and sequencing abilities. The task involves listening to, and remembering a string of digits and letters read aloud at a speed of one per second, then recalling the information by repeating the numbers in chronological order, followed by the letters in alphabetical order.

- The **PSI (Processing Speed Index)** measures speed and accuracy of visual identification, decision making and decision implementation. Performance on the PSI is related to visual scanning, visual discrimination, short-term visual memory, visuomotor coordination, and concentration. It is composed of two subtests:
  - **Cancellation:** Children scan a two-page spread of relatively small coloured pictures. The pictures include animals and objects and the child's task is to identify and cancel as many animals as possible in a given amount of time.
  - **Coding:** The child is given a worksheet where the first line contains a key in which the numbers 1 to 9 are each paired with a different symbol; his/her task is then to use this key to put in the appropriate symbols for a list of numbers between 1 and 9. The test was designed to measure speed of processing but performance is also affected by other cognitive abilities such as learning, short-term memory and concentration.

2.2.1 Performance of children with ASD: discrepancies between RPM and FSIQ (WISC-IV)

There is evidence that the evaluation of cognitive abilities in children with ASD is hard to interpret. While in typical populations there is homogeneity across and between both RPM and Wechsler FSIQ measures, some studies have found that in children with autism, these two psychometric tests provide different results (Snow et al., 1984). Studies have typically reported significantly higher RPM scores than WISC-IV FSIQ scores, in both high and low functioning children with ASD. RPM > FSIQ group performance was found in a study of Asperger syndrome (Hayashi et al. 2008) and in a study of high-functioning children with autism (Dawson et al. 2007; Soulières et al. 2009). Nader et al. (2016) specifically compared individual scores of 24 high functioning children (aged 6-16) on these two measures and found that 20% of the participants performed significantly higher on the RPM than on the FSIQ (a 50 percentile point difference or more), while no child displayed FSIQ > RPM performance. An even more marked difference between strengths and weaknesses in cognitive abilities has been found in LF children with ASD. Courchesne et al. (2015) tested 30 minimally verbal and nonverbal children with ASD (aged 6 to 12 years) both on RPM and the totality of the WISC-IV psychometric test. Results showed that while no child could complete the WISC-IV in its entirety (required in order to obtain a final FSIQ score), 87% of the group was able to complete the RPM task, with 2/3 of the group performing in the normal range.

Why do children with ASD perform so differently on RPM and WISC-IV FSIQ? This is mainly due to the nature and the structure of the tasks. As said above, while RPM is a NV task that evaluates fluid and general intelligence and minimizes the need for task instructions, culture- or experience-specific abilities, and other specific abilities (e.g., fine motor or speech skills), the Wechsler scale FSIQ involves the administration of several subtests, some of them culture-specific, which assess some specific abilities that are typically considered deficient in children with ASD. It has been reported (section 2.1) that children with ASD normally display peaks of abilities on NV tasks and valleys of performance on tasks evaluating working memory, processing speed and verbal skills, which thus impact negatively on WISC-IV FSIQ performance.

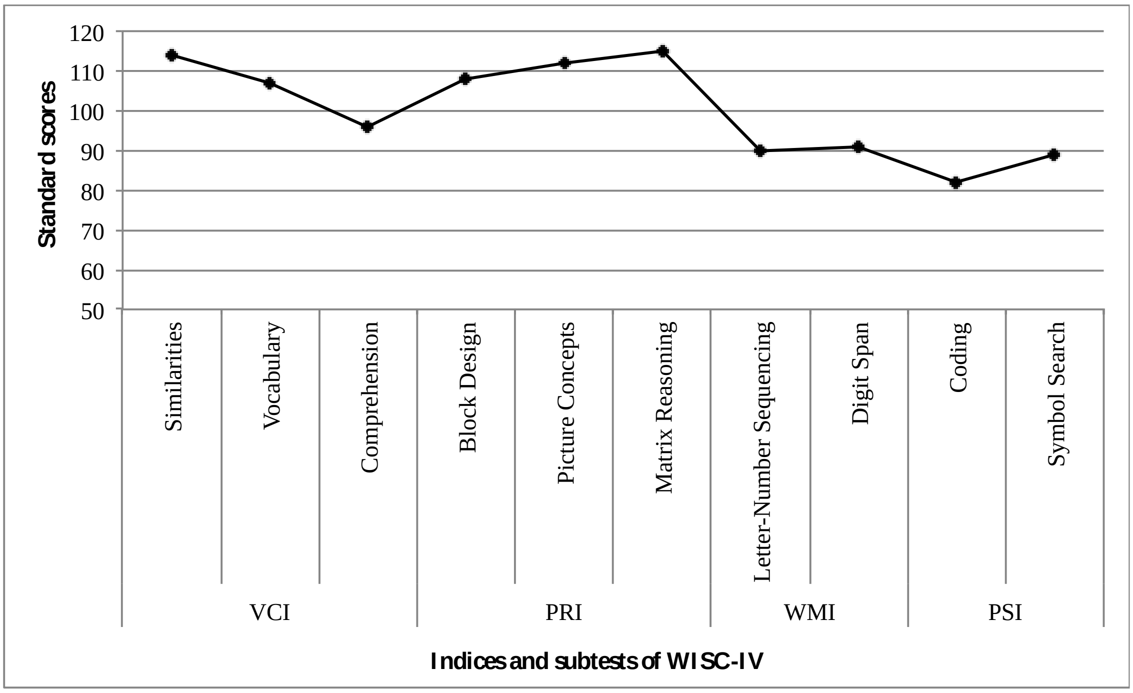
Therefore, to assess autistics' intellectual potential, RPM has been argued to be a better choice than WISC-IV FSIQ, due to the fact that it highlights the strengths of the autistic population, encompassing a wide variety of autistics, including Asperger children, HF children and even school-aged autistic children who have very little or no speech.

2.2.2 Performance of children with ASD: discrepancies between indices of the WISC-IV

Concerning the internal structure of the WISC-IV, it has been noticed that while in typical populations, testing provides similar estimates between performance on the four indices, this does not hold true for children with ASD, for whom it is not uncommon to obtain heterogeneous profiles (Barbeau et al., 2013). Children with ASD, either HF or LF, typically display large heterogeneity in results on the four indices (Oliveras-Rentas et al., 2012). Here we report the results of the study by Mayes & Calhoun (2008), who evaluated the performance of 54 HF (FSIQ > 70) children with ASD (6-14 years old) on the four indices of the WISC-IV. This study presents a paradigmatic example of this heterogeneity in a large group of children all of whom were able to perform on the totality of the subtests of the WISC-IV. The same general tendency was reported in the text manual of the WISC-IV (2003) in the section concerning populations with pathology and in several other studies evaluating intellectual abilities via the WISC-IV in groups of Asperger's children and children with HF capabilities (Stevenson & Gernsbacher, 2013 for an overview).

When assessed on all ten core subtests of the Wechsler scale, the 54 children with HF autism tested in Mayes & Calhoun, 2008, displayed a characteristic profile, illustrated in Figure 3.

FIGURE 3. COGNITIVE PROFILE OF CHILDREN WITH ASD ON WISC-IV BATTERY (DATA TAKEN FROM MAYES AND CALHOUN, 2008)



Results showed that the profile of HF children with ASD generally reflected peaks of abilities in visuospatial and perceptive domains (the highest in Matrix Reasoning subtest) and valleys in performance on the tasks evaluating attention, graphomotor skills, and processing speed (the lowest in Coding subtest). More generally, looking at the indices, as can be seen in Figure 3, a marked low performance on the PSI (processing speed and graphomotor abilities) and on the WMI (attention and working memory) was observed. On the other hand the PRI (NV, visuospatial and perceptive skills) constituted a relative peak of abilities. Further evidence of this was given by the metanalysis of Stevenson & Gernsbacher (2013), who summarized the results of nearly 40 studies of autistic participants' performance on Wechsler intelligence test batteries. On average, autistic participants performed two-thirds of a standard deviation higher on the Matrix Reasoning and the Block Design subtests than they did on other subtests, illustrating a relative autistic strength on both of these tasks. Concerning verbal abilities (VCI), HF children with ASD generally displayed good performance. However, the participants exhibited different and uneven profiles between the subtests, with the Comprehension subtest yelding significantly lower performance than the Similarities and Vocabulary subtests. These results indicate a weakness in language comprehension and social reasoning. The Comprehension task, as we said in the description above, was designed to measure social and practical understanding of a given conversation in a context by including questions, such as ‘‘What is the thing to do if you find an envelope in the street that is sealed, addressed, and has a new stamp on it?’’ and ‘‘What is the thing to do when you cut your finger?’’. Similarities and Vocabulary, instead, involve knowledge of individual words, which can be a spared domain in HF children with autism (for an overview see Walenski et al., 2006). The processing of social materials, which includes the use of language in context and the interpretation of social cues, as in the case of the Comprehension subtest, has been reported to be an area of deficit in children with ASD (Minshew et al. 2002). These findings were consistent with data reported for children with autism in the WISC-IV manual (Wechsler, 2003) and with previous studies using former versions of the Wechsler scales, the WISC-R and WISC-III (Happé, 1994; Siegel et al., 1996, a.o.).

It has been reported that LF children with ASD typically show an even more marked difference between indices of cognitive abilities and peaks and valleys of performance. Coming back to the study by Courchesne et al.(2015), which as far as we know is the only one that has looked at cognitive abilities in LF children with ASD, although no child could perform on the totality of the WISC-IV battery, six minimally verbal and nonverbal children with autism were able to complete a few subtests. One child could complete the PRI in its entirety; three completed two subtests of the PRI (Block Design and Matrix Reasoning), and two completed only one subtest of the PRI, the Matrix Reasoning. No child could complete any of the subtests of the VCI, WMI and PSI. The fact that minimally verbal LF children could complete a few subtests taken from the PRI task reinforces the hypothesis that NV abilities are actually a strength in the autistic population, independently from their cognitive level.

Why do children with ASD perform so differently on indices of the WISC-IV? Similarly to what we concluded for the RPM/WISC-IV FSIQ discrepancy, this is mainly due to the nature and the structure of the indices of the Wechsler scale. While the PRI is composed of NV tests (which are relatively culturally-fair and language free) evaluating fluid and general intelligence and visuo-spatial skills, the other three indices of WISC-IV

(VCI, WMI and PSI) include tasks evaluating working memory, processing speed and verbal skills, which are known to normally constitute areas of deficits for children with ASD all across the spectrum. Therefore, to assess autistics' potential, the PRI has been suggested to be a better choice than all the other indices of the WISC-IV, since it allows to have a full picture of children on the spectrum, encompassing a wide variety of autistics, which includes Asperger children, HF children and even school-aged autistic children who have very little or no speech. We note that the WISC-IV test manual clearly states that when a child displays a significant difference between indices, the FSIQ should not be considered to be a reliable measure of general intelligence, and suggests that the subtests be treated separately in order to highlight an individual's strengths and weaknesses (Wechsler, 2003).

To sum up:

RPM and the PRI of WISC-IV have each been argued to be the most suitable tools for evaluating autistic intelligence. Why is this so?

First of all, general intelligence, as evaluated in the WISC-IV test via the FSIQ score, is not wholly appropriate for individuals on the autism spectrum, whose performance is markedly different depending on the cognitive domain involved. An evaluation of cognitive abilities through the FSIQ score does not differentiate between peaks (NV intelligence) and valleys (verbal intelligence, working memory and speed processing) of abilities, and it arguably therefore underestimates the intelligence of children with autism (Nader et al., 2016). The nature and the structure of both RPM and the PRI make it possible for the "real" abilities of children with ASD to be highlighted.

Secondly, both RPM and the PRI can be used as psychometric tools for evaluating children from the entire spectrum (independently of their level of cognitive ability). While HF children typically succeed in performing a wide variety of tasks evaluating different cognitive abilities, this does not hold true for children with LF cognitive abilities. Standardized intelligence tests such as the WISC-IV are often of limited usefulness in the assessment of children with severe intellectual disability, who are sometimes unable to stay on task for the lengthy administration of the test, have trouble handling its heavy reliance on language skills, or lack the ability to be motivated (Turner et al., 2006). Use of a general measure of intelligence which is not adapted to the needs and capacities of both HF and LF children with ASD would lead to all LF children being unable to achieve a basal score of

FSIQ. In order to obtain a valid measure of cognitive ability for children with severe intellectual deficiency, testing procedures must accommodate their profound deficits in communication, attention and social skills. In LF children, who are unable to perform on the totality of the psychometric measures, RPM and the PRI may be the only tasks that reflect the enhanced perceptual performance of these children (Samson et al., 2012). Nonverbal tasks such as RPM and the PRI seem to be, then, natural choices for evaluation of cognitive abilities in children with ASD, considering that perceptual processing, often manifested as enhanced visuospatial and abstract reasoning abilities central to the concepts of fluid and general intelligence, could be defined as an associated feature of the autistic phenotype, both in HF and LF children with ASD.

2.2.3 RPM and PRI: most appropriate tools?

Going back to the purpose of the present study, we need to directly address the question of whether RPM and PRI satisfy all three criteria we put forth for our definition of "most appropriate tools" for children with ASD (see Section 2.1).

Our first criterion was that such an appropriate tool should be easily available and that it can be easily used both in research and clinical practice. Both RPM and the PRI (WISC-IV) are the most frequently used tools for evaluating human intelligence and cognitive abilities in children with ASD. These tools have been widely standardized across the world.[7] This facilitates both their use in research and clinical practice, and in inter-study comparisons. Moreover, both RPM and Wechsler scales are frequently used to assess cognitive abilities in the general population. This facilitates comparison with control groups.

The second criterion should be that such tools should stress the “real” cognitive abilities of individuals on the autism spectrum, without any confounding effects of other variables (much like what repetition tasks did for structural language abilities). In this sense, both RPM and the PRI stress the real abilities of individuals on the spectrum (enhancing

[7] The WISC has been translated or adapted in many languages, and norms have been established for a number of languages, including Spanish, Portuguese (Brazil and Portugal), Arabic, Norwegian, Swedish, Finnish, Czech, Croatian, French (France and Canada), German (Germany, Austria and Switzerland), English (United States, Canada, United Kingdom, Australia), Welsh, Dutch, Japanese, Chinese (Hong Kong), Korean (South Korea), Greek, Romanian, Slovenian, Indian and Italian. Separate norms are established with each translation (Wechsler, 2003). RPM is the test that had the greatest number of adaptation/standardizations in the world, although the exact number does not seem to be available (Raven, 2000).

both strengths and weaknesses of this clinical population). The chosen tools are in fact suitable for an evaluation of cognitive abilities on the entire spectrum. As said before, children with ASD typically excel on visuo-spatial tests and measures of abstract reasoning. This statement holds true for the totality of the spectrum and it has been confirmed by previous work that relied on previous diagnostic criteria of the DSM-IV, which provided a division between Autistic Disorder, AS and PDD-NOS (Courchesne et al., 2015; Mayes & Calhoun, 2008; Nader et al., 2016; Oliveras-Rentas et al., 2012; Soulières et al., 2009; Stevenson & Gernsbacher, 2013). This means that psychometric evaluation via tasks of perceptive and NV abilities (as RPM and the PRI) should guarantee, on one hand, uniform evaluation of cognitive abilities through the spectrum and, on the other hand, it should assure reliable comparison between individuals on the totality of the spectrum. This is not true of tasks evaluating other aspects of cognition, such as verbal intelligence, working memory skills, speed processing, due to the great heterogeneity of performance displayed by children with ASD.

Thirdly, the appropriate tools should be at most remotely related to language abilities. Since the purpose of the present work is to detect profiles of linguistic and cognitive abilities in children with ASD cognitive scores should be as nonverbal as possible. Both RPM and the PRI seem to meet this criterion, being NV tasks. The use of a tool or even a battery evaluating multiple domains of intelligence, including verbal abilities, may particularly affect the performance of those children displaying impaired language abilities. The use of tasks requiring some form of verbal abilities (which may range from the evaluation of language abilities, as in the case of the VCI of WISC-IV, to a task that depends on the capabilities of an individual to comprehend verbal stimuli and/or verbalize the answer, as in the case of the Digit Span subtest) raises the complementary question of the extent to which low cognitive performance could be the result of low language abilities, in some children with ASD.

Since both RPM and the PRI satisfy all three criteria we have set for the present study, they would appear to be appropriate measures to select. However, before firmly drawing the conclusion that RPM and the PRI are the best tools for evaluating cognitive abilities in relation to language abilities in children with ASD, we need to fill two gaps in the literature regarding their appropriateness/usefulness. First, although RPM has been indicated as the most suitable test for evaluating intelligence in research on ASD, and the PRI has been pinpointed as the most appropriate index of the WISC-IV for picturing autistic intelligence in the clinical domain (Nader et al. 2016), no study has compared these two measures. Performance of children with ASD on these two tests need to be compared in order to verify whether these two tools actually evaluate similar NV abilities or not. Second, while RPM is often described as a ‘‘non-verbal’’ test, it has been argued that in the non-autistic population, verbal abilities may be crucial in improving performance on this task (Fox & Charness, 2010; Raven's manual 1998). Furthermore, within the PRI, a few studies have reported an additional valley of abilities on the Picture Concepts subtest when compared to the Matrix Reasoning and the Block Design subtests (Nader et al., 2015; Oliveras-Rentas et al., 2012). As suggested by (Houskeeper, 2011), children with ASD may display more difficulties on this task because, although it is included in the perceptual index of the WISC-IV and is defined as a task evaluating fluid and general intelligence, it strongly relies on language abilities since children are asked to verbalize their answers (see above). This differs in this regard from Block Design (a visuospatial task) and Matrix Reasoning (a pure fluid reasoning test). To our knowledge, no study has directly verified the relation between linguistic abilities and performance on both RPM and the subtests of the PRI. Since the present work sought to use of tasks that are nonverbal in nature (minimal or even no need for task instructions or speech skills), we will verify the possible relations between linguistic abilities on the one hand, and RPM and each of the subtests of the PRI on the other hand, in children with ASD.

2.2.4 Conclusions for the evaluation of autistic intelligence

The heterogeneity of performance that children with ASD may display between different psychometric tests (RPM vs. FSIQ), between different domains of cognitive abilities (e.g., PRI vs. VCI, WMI and PSI, in the WISC-IV) and sometimes between tasks evaluating the same domain of cognitive abilities (Picture Concepts vs. Block Design, in the PRI) have obvious implications for determining the prevalence of intellectual disability in autism and the detailed nature of the autistic cognitive phenotype. As abilities of children with ASD are routinely assessed using comparisons with intelligence-matched controls, the choice of an appropriate intelligence measure is a crucial procedural decision for both high-functioning and low-functioning individuals. Peaks and valleys may differentially influence averaged level measures according to the instruments being used. Over- and/or under-estimation of cognitive abilities are frequent when peaks and valleys are not well separated among cognitive performance in children with ASD.

Thus, a number of questions arise regarding how studies that have evaluated language abilities in children with ASD have separated children into HF and LF subgroups. We will see in the next section (2.3) that the studies on language in ASD we reported in Chapter 1 did not systematically take into account the strengths and the weaknesses of the ASD population when they reported on cognitive levels in the groups of children under examination. We will see that intellectual abilities were evaluated via different tests and instruments; furthermore, the results were pooled together in broad bands of intellectual levels whose boundaries vary from study to study. The heterogeneity of psychometric tools used for the evaluation of cognitive abilities in children with ASD has hampered clear conclusions on the genuine abilities of children with ASD and at the same time has made inter-study and inter-group comparisons very difficult. As a consequence, the reported cognitive abilities in these studies should be interpreted with caution. The description of autistic intelligence we reported in section 2.2 was necessary to comprehend which are the measures most likely to appropriatley assess cognitive abilities in relation to structural language abilities in children with ASD. With these *a priori* in mind, we turn now to the choice of psychometric tools used in the studies evaluating structural language abilities in children with ASD reported on in Chapter 1.

2.3 How intellectual abilities have been evaluated in studies assessing language abilities in children with ASD: introduction

In Chapter 1 we decided, for the sake of clarity, to use the labels HF and LF in accordance with the most frequent cognitive ability labels used in the literature. Essentially, we provided the HF label to all participants who performed within the norms on any given psychometric test, and the LF label to all the participants who performed below the norms. In other words, the division was made independently of the test employed for evaluating cognitive abilities and on the basis of the cut-off given in each study for normal and impaired performance. This decision was made for the purposes of Chapter 1, which concentrated on the topic of structural language abilities. We have just seen, however, that results on cognitive abilities of children with ASD are in fact hard to interpret. Studies (section 2.2) have shown that specific tools adapted to the specific needs of individuals on the autism spectrum should be used for the evaluation of their cognitive abilities. We suggested that when cognitive abilities are directly compared to linguistic abilities, the nonverbal nature of the evaluation of cognitive abilities needs to be controlled. In this section we will report on how studies investigating linguistic abilities in children with ASD (the studies reported in Chapter 1) have evaluated cognitive abilities in their population samples. We will limit our review to the studies that have provided clear information about the psychometric tool used for assessing intellectual abilities and the mean rate of performance of the ASD group (5 studies will be excluded from the present review). We will summarize, in a recapitulatory table (Table 4), the methodological characteristics of cognitive evaluations adopted by these studies (e.g. the psychometric tools, the criteria for separating children into HF and LF profiles, and the cut-offs and mean rates of performance when provided). We will see that much like the ASD-LI/ASD-LN distinction, the difference between HF and LF in previous work rests on the use of different tools and different criteria across studies.

2.3.1 Evaluation of intellectual abilities in studies assessing language skills in children with ASD: description, limitations and hypothesis for the present study

Table 4 presents a summary of all of the psychometric tools and the cut-off criteria used to evaluate cognitive abilities in studies assessing structural language in children with ASD (Chapter 1). Based on this review, we will identify two potential methodological stumbling blocks, which we believe constitute major limitations for the question of language profiles in ASD: 1) The heterogeneity of psychometric tools and cognitive domains, and the criteria used for separating children into groups of HF and LF capabilities have prevented clear conclusions on the real abilities of children with ASD and at the same time have made inter-study and inter-group comparisons (almost) impossible and 2) if we apply our list of criteria for the choice of reliable psychometric tools to this literature review, we can see that these criteria often collide with the specific methodological choices made by previous studies.

Table 4. Overview of the psychometric tools and criteria used to define children with ASD with HF and LF cognitive abilities of studies listed in chapter 1

	Study	Psychometric test	Criteria for division between HF and LF subgroups (Cut-off)	Mean (SD); range of cognitive scores (when given)	HF and/or LF population
1	Eigsti & Bennetto (2009)	WISC-III FSIQ	≥ 85 standard score	119 (14); 91-138	HF
2	Shriberg et al. (2001)	WISC-III both VIQ and PIQ		VIQ = 106 (27); PIQ = 86 (27)	HF
3	Zhou et al. (2015)	WYCSI (Wechsler Chinese test) VCI	> 80 standard score	101 (12)	HF
4	Bartak 1975	WISC NVIQ	≥ 70 standard score	96 (13)	HF
5	Boucher 1976	WISC NVIQ	≥ 70 standard score	70-110	HF
6	Loucas et al. (2008)	WISC-III PIQ	≥ 80 standard score		HF
7	Bishop (2004)	WISC-III PIQ	≥ 70 standard score	82 (25)	HF
8	Williams et al. (2013)	WISC-IV PRI	> 80 standard score	92 (12)	HF
9	Ellis Weismer et al. (2017)	WISC-IV PRI	> 80 standard score	107 (13)	HF
10	Hill et al. (2015)	WISC-IV PRI or WIPPSI-III PIQ	≥ 80 standard score	112	HF
11	Botting & Conti Ramsden (2003)	WISC-III PIQ (Block and Picture Completion)	≥ 70 standard score	90; 76-106	HF
12	Riches et al. (2010)	WISC-III PIQ (Block and Picture Arrangement)[a]	≥ 80 standard score	105 (2); 90-130	HF
13	Riches et al. (2011)	WISC-III PIQ (Block and Picture Arrangement)	≥ 80 standard score	105 (2); 90-130	HF
14	Diehl et al. (2015)	WASI FSIQ	≥ 80 standard score	113 (16); 89-136	HF

15	Whitehouse et al. (2008)	WASI PIQ	≥ 80 standard score	100 (13); 80-137	HF
16	Norbury et al. (2010)	WASI Matrix Reasoning subtest	> 85 standard score	100 (10)	HF
17	Taylor et al. (2014)	WASI Matrix Reasoning subtest		105 (2)	HF
18	Cleland et al. (2010)	RPM		97 (15)	HF
19	Su et al. (2014)	RPM	≥ 80 standard score	115 (14)	HF
20	Schaeffer (2017)	RPM		64th percentile	HF
21	Terzi et al. (2012)	RPM	≥ 80 standard score		HF
22	Terzi et al. (2017)	RPM	≥ 80 standard score		HF
23	Brynskov et al. (2016)	RPM	≥ 80 standard score	106 (19); 80-140	HF
24	Harper-Hill et al. (2013)	RPM	≥ 80 standard score	100 (3)	HF
25	Naigles et al. (2011)	MSEL Visuospatial abilities	developmental scale comparable with TD age peers (age in months)	75 (6)	HF
26	Goodwin et al. (2012)	MSEL Visuospatial abilities	developmental scale comparable with TD age peers (age in months)	75 (6)	HF
27	Tovar et al. (2015)	MSEL Visuospatial abilities	developmental scale comparable with TD age peers (age in months)	79 (15)	HF
28	Jyotishi et al. (2017)	MSEL Visuospatial abilities	developmental scale comparable with TD age peers (age in months)	84 (5)	HF

29	Nadig et al. (2017)	MSEL Receptive and Expressive Language subscales	developmental scale comparable with TD age peers (raw scores)		HF
30	Tek et al. (2014)	MSEL Expressive Language subscale		18 (8); 9-33 raw scores	HF
31	Janke & Perovic (2015)	KBIT Matrix subtest	≥ 80 standard score	113 (21); 82-154	HF
32	Khetrapal et al. (2017)	KBIT Matrix subtest	≥ 70 standard score	94 (12); 74-121	HF
33	Tager Flusberg 2006	DAS NVIQ		83	HF
34	Tager Flusberg 2015	DAS NVIQ		FSIQ: 102 (18) VIQ: 95 (16) NVIQ: 104 (20)	HF
35	Waterhouse & Fein (1982)	Matching figures test (Kagan's preschool) & Frosting test of visual perception		7.5 (5)	HF
36	Lloyd 2006	Assessed by an educational psychologist (no other information)			HF
37	Park et al. (2012)	WISC-III PIQ		89 (18); 59-116	HF and LF
38	Jensen de Lopez et al. (2018)	WISC-IV Matrix Reasoning subtest	≥ 85 standard score	70-105	HF and LF
39	Tuller et al. (2017)	WISC-IV PRI or WPPSI-III PIQ and RPM	≥ 80 standard score	40-110	HF and LF
40	Durrleman et Delage (2016)	RPM	≥ 80 standard score		HF and LF

41	Zebib et al. (2013)	RPM	≥ 80 standard score		HF and LF
42	Prévost et al. (2017)	RPM	≥ 80 standard score		HF and LF
43	Durrleman et al. (2017)	RPM	≥ 80 standard score		HF and LF
44	Weismer et al. (2011)	Bayley Scales for Infant Development		85 (10); 57-105	HF and LF
45	Rapin et al. (2009)	Stanford-Binet Intelligence Scale Verbal and Nonverbal IQ			HF and LF
46	Eigsti et al. (2007)	Standford Binet Nonverbal IQ		80 (15) 49-115	HF and LF
47	Perovic et al. (2013)	KBIT Matrix subtest		66 (22); 40-105	HF and LF
48	Modyanova et al. (2017)	KBIT Matrix subtest		74 (22); 40-145	HF and LF
49	Kjelgaard & Tager-Flusberg (2001)	Differential abilities scales FSIQ, VIQ and NVIQ		FSIQ: 68 (24) 25-141 VIQ: 76 (19) 51-133 NVIQ: 85(20) 43-153	HF and LF
50	Roberts et al. (2004)	Differential abilities scales FSIQ, VIQ and NVIQ		FSIQ: 68 (24) 40-149 VIQ: 76 (19) 51-133 NVIQ: 85(20) 43-153	HF and LF
51	McClery et al. (2006)	Bayley Scales for Infant Development			LF

Note: Studies are listed in the following order: studies that included only HF children; studies that included both HF and LF children and studies that included only LF children. Empty cells indicate information that was not provided.

[a] The Picture Arrangement subtest includes 4 to 6 pictures that are part of a story, presented in a mixed up order to participants. The child needs to rearrange them into a logical story sequence, on the basis of his/her knowledge of the word and capacity of integrating multiple information (Wechsler, 1991).

From the literature review presented in Table 4, we can draw three main conclusions. First, the psychometric tools employed for evaluating cognitive abilities of children with ASD were very heterogeneous and all included different indexes and different subtests targeting different abilities. In our literature review we detected eight tasks: Wechsler scales (in the following versions WISC, WISC-III, WISC-IV, WASI, WPPSI); Raven's Progressive Matrices (RPM); Mullen Scales for Early Development (MSEL); The Kaufman Brief Intelligence Test (KBIT); Differential Abilities Scales (DAS); Bayley Scales for Infant Development; Stanford-Binet Intelligence Scale; Kagan's Preschool test. Although Wechsler scales and RPM were the most frequently adopted tools (31/51 studies), other batteries were used as well. These other psychometric tools, much like WISC scale, included different indexes and different subtests targeting different abilities. Several studies compared performance of TD children across test batteries - KBIT and WISC (Chin et al., 2001), DAS and WISC (Dumont et al., 1996), Bayley scale and WISC (Månsson et al., 2018), suggesting that both results on FSIQ scores and indexes evaluating different abilities reached 95% limits of agreement. Only two studies compared performance of children with ASD on the MSEL and the DAS battery (Bishop et al., 2011) and on the WISC and the Standford-Binet battery (Mayes & Calhoun, 2003), and in both cases, profiles of cognitive abilities were generally consistent on both tests. The use of different tools (which may be due to different reasons, e.g. age of the participants, use of particular IQ tests in the centres where the children were recruited, specific research agendas, etc.) does not constitute *per se* a limitation. What may really jeopardise the reliability of inter-group comparisons is that the same types of measures (domain of abilities and cut-off criteria for intellectual impairment) were not used across the board in studies on language in ASD.

Among the studies that employed the same tests, in fact, we identified the use of a wide variety of measures for dividing children into HF and LF profiles: for the Wechsler scales some studies employed the Full Scale IQ score, while others used the scores drawn from the Verbal Index, the Perceptual Index, or some specific subtests of the Perceptual Index (Block Design, Matrix Reasoning, Picture Completion, Picture Arrangement); for the MSEL some studies used the Visuospatial index and others the Expressive/Receptive Language index; for the Standford-Binet Test some studies employed the Verbal index score, while others used the Nonverbal index score. As seen earlier, the use of FSIQ scores or Verbal indexes of cognitive abilities does not provide the same results as the use of Nonverbal indexes in children with ASD, creating an imbalance between weaknesses and strengths in the evaluations.

In addition, the criteria for separating children into profiles of HF and LF abilities vary consistently across studies. For Wechsler scales, and more generally for standardized psychometric batteries, the cut-off criteria fluctuates between ≥ 70 and ≥ 85 standard scores. Five studies considered children with autism having HF profiles when their performance on the psychometric tests was ≥ 70 standard score, one at ≥ 75 standard scores, two at > 80 standard score, 17 at ≥ 80 standard score, and three at ≥ 85 standard score. Moreover, even among the studies that used the same tests, the cut-offs used for indicating intellectual impairment varied consistently from one study to the other, e.g. Loucas et al. (2008) selected a cut-off of < 80 standard score for the NVIQ index of the WISC scale, while Bishop (2004) used a threshold of < 70 standard score. Looking in detail at the Wechsler scale manual, the traditional cut-off score for intellectual disability (ID) is established at < 2 SD below the mean (i.e. a standard score of < 70). Here we report the descriptive classification of IQ provided in the WISC-IV manual.

TABLE 5. DESCRIPTIVE CLASSIFICATION OF IQ USING THE SCALE OF WISC-IV MANUAL

<table>
<tr><th>Classification</th><th>Standard scores</th><th></th></tr>
<tr><td>Very superior</td><td>≥ 130</td><td></td></tr>
<tr><td>Superior</td><td>120-129</td><td></td></tr>
<tr><td>High average</td><td>110-119</td><td></td></tr>
<tr><td>Average</td><td>90-109</td><td></td></tr>
<tr><td>Low average</td><td>80-89</td><td></td></tr>
<tr><td>Borderline</td><td>70-79</td><td></td></tr>
<tr><td>Mild ID</td><td>55-69</td><td rowspan="4">Extremely Low IQ</td></tr>
<tr><td>Moderate ID</td><td>40-54</td></tr>
<tr><td>Severe ID</td><td>25-39</td></tr>
<tr><td>Profound ID</td><td>< 25</td></tr>
</table>

Looking back at Table 4, we can see that in research, results were pooled together in broad bands of intellectual levels which do not share the same cut-offs across studies. Essentially, studies differed on the status of the *borderline* profile. Some studies considered children with borderline IQ (standard score from 70 to 79) as HF children, while others considered them as LF children, depending on the cut-off they applied. The same consideration holds true for studies using the MSEL scale (which separated children on the basis of mental age and studies assessing cognitive abilities via the KBIT scale). For RPM the cut-off for intellectual impairment was generally established at < 80 standard score, which means that typically, in research, studies using Raven's matrices as measures of NVIQ considered children with borderline abilities as having a LF profile. The heterogeneity of criteria used for establishing a cut-off of spared and impaired cognitive abilities prevents clear conclusions on the intellectual abilities of children with ASD and at the same time has made inter-study and inter-group comparisons unreliable.

Finally, we can see that some specific methodological choices made in previous studies cannot be adapted for the purpose of our study. Following the third criterion of our list in section 2.1, the use of FSIQ scores or verbal tasks to represent cognitive abilities in children with ASD, as some of the studies listed in Table 4 did, may lead to inappropriate subgroupings, not considering that the choice of cognitive tools involving language abilities may have led to potential confounds. An inter-study/inter-group comparison with our study will be limited, in the discussion of the present work, to the studies that employed RPM and/or the PRI (or at least one of the subtests of the PRI) of WISC-IV (19/51 studies).

2.4 General conclusions and direction of research for the present study

Research on intellectual profiles in children with ASD suggests that when a psychometric tool is used for the evaluation of cognitive abilities in children with ASD, we need to keep in mind that children on the autism spectrum might show considerable discrepancies between their performance on different intelligence measures and on different domains of cognitive abilities, differences not found in the general population. This result has not been systematically incorporated into studies on language in ASD in the way they identify subgroups of children on the spectrum.

In this chapter, we reported on current findings in the literature that argue for both Raven's Progressive Matrices (RPM), and the Perceptual Reasoning Index of the 4th edition of the Wechsler Intelligence Scale for Children (WISC-IV) as the best-suited measures for evaluating cognitive abilities in children with ASD. Essentially they can be easily used both in research and in clinical practice; they have been argued to reveal a complete picture of autistic intelligence and they allow for evaluation across the spectrum. Moreover, and this is important for the purpose of the present study, being NV tasks, they should not be linked to verbal abilities. These tasks should therefore provide the foundations for meaningful exploration of the relation between linguistic abilities and cognitive abilities in ASD. A crucial outstanding issue for the use of these nonverbal tasks is the possible link between linguistic abilities and RPM and PRI subtests, which has not been specifically addressed in the literature. Likewise, nothing is known about potential discrepancies between RPM and PRI subtests in the performance of children with ASD.

Finally, we can conclude that the studies that concentrate on the description of language abilities in autism should pay more attention to the choice of psychometric measures of cognitive abilities, especially when they are put in relation with linguistic abilities.

Chapter III
Autism severity and developmental factors

3.1 Introduction

Before moving to the experimental part of the present work, we need to briefly consider the possible relation between linguistic/cognitive abilities and both severity of autism symptoms and developmental factors. There have been several controversies concerning whether generalized measures of severity of autism symptoms might be related to the performance of children in this population on formal aspects of language and/or to intellectual disability, and whether developmental factors such as age of first word and age of first sentence may have an impact in predicting outcomes of language and cognitive abilities in this population (Charman et al., 2011; Siller & Sigman, 2008).

In the next section we will report on the few studies that have investigated the relation between severity of autism symptomatology and linguistic/cognitive abilities. We will see that not only have reported results been mixed, but there has been general misunderstanding as to which tools properly measure autism severity, leading to a greater confound in the literature. In the second part of the chapter we will focus on developmental factors that have been reported to be good predictors of both structural language abilities and cognitive abilities in children with ASD, namely age of first word and age of first sentence.

3.2 Autism severity and linguistic/cognitive abilities

Very few studies have investigated extra-linguistic and extra-cognitive factors that may be related to formal language impairment and intellectual impairment in ASD. Among these, some studies have found significant correlations between severity of autism and language abilities (Eigsti & Bennetto, 2009; Harper-Hill et al., 2013; Modyanova et al., 2017; Tek et al., 2014; Whitehouse et al., 2008), while others have found no such significant correlations (Hill et al., 2015; Loucas et al., 2008). Similarly, studies investigating cognitive abilities in children with ASD have shown contrasting results, between those suggesting that cognitive level of children with autism was strongly related to severity of autism symptoms (Luyster et al., 2008) and those that did not find any clear correlation (Charman et al., 2011; Joseph et al., 2002). Why is there this heterogeneity of results? To answer this question it is necessary to consider how autism severity was assessed in these studies.

Severity of autism symptoms has often been defined in terms of language ability, intellectual functioning or the presence of problematic behaviours, all of which contribute significantly to an individual's ability to function in social situations (Weitlauf et al., 2014). However, these are not necessarily core features of autism as the disorder is currently being conceptualized. If severity of autism is defined as the degree or intensity of autism itself, we need to direct our attention to the core features of autism (the two dimensions of the DSM-5), namely social and communicative impairments, and repetitive or restricted interests and behaviours. On this definition, severity has very often been estimated via direct measures of ASD assessment, such as the Autism Diagnostic Interview-Revised, ADI-R (Le Couteur et al., 2003) and the Autism Diagnostic Observation Schedule, ADOS (Lord et al., 2003). Along with diagnostic manuals, the DSM-5 and the ICD-10, both ADI-R and ADOS diagnostic algorithms are used to yield a classification of "autism" or "nonautism". These tools exhibit good to excellent predictive validity, especially when used in combination with each other (Risi et al., 2006). However both ADI-R and ADOS were conceived as diagnostic tools and not as severity metrics for core autism symptoms (Hus & Lord, 2013).

The ADI-R as described in the clinical manual is a standardized, semi-structured parent interview that is administered face-to-face by a trained clinician in approximately 2-3 hours. It includes 93 items focusing on Reciprocal Social Interactions, Language/ Communication Restricted, Repetitive Behaviours and Interests and Early Development. A total score is calculated for each of the interview's content areas. Then an algorithm is applied for final diagnosis.

The ADOS is a semi-structured assessment of communication, social interaction, and play (or imaginative use of materials) for individuals suspected of having autism. The ADOS consists of four modules, each of which is appropriate for children of differing developmental and language levels, ranging from nonverbal to verbally-fluent. Module 1 (no words) is intended for nonverbal children. Module 1 (few words) is intended for individuals who do not consistently use *phrase speech* (defined as non-echoed, three-word utterances that sometimes involve a verb and that are spontaneous, meaningful word combinations). Module 2 is intended for individuals with some phrase speech who are not verbally fluent. Module 3 is intended for verbally fluent children for whom playing with toys is age-appropriate (usually up to 12 - 16 years of age). *Verbal fluency* is broadly defined as having the expressive language of a typical four-year-old child: producing a range of sentence types and grammatical forms, using language to provide information about events out of the context of the ADOS, and producing some logical connections within sentences (e.g., "but" or "though").

Gotham and colleagues (2009) were the first to pinpoint the fact that ADOS classification thresholds (Autism, Autism Spectrum, Non-Spectrum) and ADOS/ADI-R raw totals have been used as additional proxies for autism severity. While it is true that higher ADI-R and ADOS scores indicate a greater number of items representing core deficits associated with ASD, raw and total scores were not normalized for this purpose. Moreover, it has been demonstrated that both these tools' global scores strongly correlate with individuals' linguistic level and IQ, which reduces the focus on autism core symptoms (Mehling & Tassé, 2016) and more importantly, would make them unfit for the purposes of the present study. Gotham et al. tried to overcome this problem by introducing a calibrated severity score for the ADOS. As suggested by de Bildt et al., (2011) "developing these calibrated severity scores was inspired by the need in clinical practice and research for describing the severity of the behaviour of children with ASD referring to the core symptoms in the autism spectrum." Standardized *calibrated severity scores* are more uniformly distributed when compared to raw scores across age and language groups, and less influenced by participants' characteristics (e.g. age differences between children administered different modules, cognitive level and linguistic abilities). Moreover, unlike raw scores, they provide a relative metric of the severity of autism-specific symptoms. The scores range from 1 to 10 and classify the child as not displaying typical autistic symptomatology (< 4), or showing typical traits of individuals on the spectrum (≥ 4). In clinical practice and research, the calibrated severity measure has been valuable for indicating the severity of specific ASD behaviour, comparing ADOS assessments across modules, investigating the relationship between severity in ASD and levels of cognitive functioning, and investigating the relationship between severity in ASD and levels of linguistic abilities (Gotham et al., 2012).

In the studies we reported on earlier and looking at the criteria used to retrieve severity of autism scores, while some used scores specifically calculated to reflect individuals' levels of autism severity, e.g. ADOS calibrated severity scores (Charman et al., 2011; Hill et al. 2015; Loucas et al., 2008; Whitehouse et al., 2008), the majority used global scores derived from diagnostic tools, e.g. the ADOS summary scores (Eigsti & Bennetto, 2009; Harper-Hill et al, 2013; Luyster et al., 2008; Tek et al., 2014) or both the ADI-R and the ADOS summary scores (Joseph et al., 2002; Modyanova et al., 2017).

This use of both summary scores of diagnostic tools and severity scores for referring to severity of autism symptoms prevents any clear conclusions on the relationship between structural language abilities and/or cognitive abilities and severity of autism symptoms in children with ASD. This methodological stumbling block makes it necessary to differentiate between tools and scores that should be used for diagnostic purposes and tools and scores that should be used for evaluating severity of autism symptoms. For the purposes of the current study, we sought severity of autism scores that were as far as possible from linguistic abilities and cognitive abilities.

A second methodological problem is represented by the applicability of ADOS and ADI-R summary scores in a developmental perspective. Both of these tools' scores are derived from the diagnostic assessment of a child, which typically takes place once and at a very early age (3 years old, for children that are detected early). This means that the scores drawn from ADOS and ADI-R do not account for the possible changes in developmental trajectory (in particular, with respect to potential changes in severity of autism symptoms) of a child with ASD. It is well known that as a child with ASD develops, he/she often moves through age-, language- and cognitive- levels. Research on developmental trajectories in ASD has focused on the stability of categorical diagnoses, verbal and cognitive outcomes, and symptom domain change over time. Gotham et al., (2012) examined longitudinal trajectories of ASD severity from early childhood to early adolescence in 345 individuals with ASD and found that, although overall the majority of children remained stable in their ASD severity scores over an 8–12 year period, there was evidence for improving, plateauing (i.e., developmental slowing) or even worsening (i.e., loss) trajectories of skills (cognitive, linguistic, etc.) in subgroups of children with ASD (see also Landa et al., 2007; Ozonoff et al., 2008). These studies support the notion that determination and specification of ASD severity at time of diagnosis, as well as re-evaluation, is valuable as it may convey important information about symptom course and prognosis. For these reasons some authors have pinpointed the necessity of using tools that can both evaluate the rate of severity of autism symptoms in a child with ASD at a given time and detect changes in symptom severity (Guthrie et al., 2013).

Within the growing body of literature on the trajectory of ASD-specific symptom expression, severity has most often been quantified with scores from the ADOS calibrated severity scores (which we described above) used as a baseline, and through evaluation scales of autism severity, among which the Childhood Autism Rating Scale, CARS (Schopler et al., 2002) is the most widely used. The CARS is a 15-item behavioural rating scale developed to identify children with autism and to categorize their behaviours from mild to moderate to severe. The 15 items in the scale are: relating to people; imitative behaviour;

emotional response; body use; object use; adaptation to change; visual response; listening response; perceptive response; fear or anxiety; verbal communication; non-verbal communication; activity level; level and consistency of intellective relations; general impressions. The examiner assigns a score of 1 to 4 for each item: 1 indicates behaviour appropriate for age level, while 4 indicates severe deviance with respect to normal behaviour for age level. The scores for the single items are added together into a total score, which classifies the child as not displaying typical autistic symptomatology (below 30), mild or moderately autistic (30–36.5) or severely autistic (above 36.5).

For the purpose of the present study, we also use another scale of autism severity, developed by the researchers at the Regional University Hospital Centre in Tours, which will serve as a complementary measure of symptom severity in children with ASD, the Echelle d'Evaluation des Comportements Autistiques (‘the autistic behaviour evaluation scale’), ECA-R (Lelord & Barthélémy, 2003). Since children at the Tours Centre are normally assessed with this tool twice a year, this provided a recent, reliable evaluation of autism symptomatology (normally not more than six months prior to the administration of our research protocol). The ECA-R is a behavioural scale specifically developed to evaluate severity of autistic behaviours. The ECA-R assesses the frequency of autistic behaviours in 29 domains, which include items that reflect aspects typically affected in children with ASD such as aloneness, incapacity to interact with people, difficulties to engage in spontaneous activities and spontaneous turn-talking, stereotyped sensory-motor activities, intolerance to changes, a.o. The examiner assigns a score of 0 to 4 for each item: 0 indicates absence of the autistic behaviour, while 4 indicates the constant presence of the autistic behaviour. The scores for the single items are added together into a total score. No cut-off criteria are assigned, because the scale was developed with the precise intent to follow and evaluate the development of autistic behaviours of an individual through the course of time.

A final consideration is needed: even if these developmental scales purported not to be related to linguistic and cognitive abilities in children with ASD, no study has directly verified these hypotheses.

To sum up:

Since one of the purposes of the present study is to verify the possible relations between structural language abilities and NV abilities with severity of autism symptoms, we decided not to use ADI-R and ADOS global scores as measures of autism symptomatology. We retrieved data for all three severity of autism measures we listed above: ADOS severity scores (which we used as baseline, as suggested in the study by Gotham et al. 2012), CARS and ECA-R. These measures should ensure a more reliable evaluation (not as related to language abilities and cognitive abilities) of severity of autism symptoms of children with ASD. We will see in Chapter 5, how we employed each one of these scales in the analysis of our results.

3.3 Developmental factors

Two developmental factors have previously been reported as good predictors of language abilities and/or cognitive abilities in children with ASD: age of first word and age of first sentence. A large body of evidence suggests that early language acquisition predicts later functional and developmental outcomes (Anderson et al., 2007; Bennett et al., 2008; Loucas et al., 2008). Age of first word has been suggested as a strong predictor for both language (Kover et al., 2016) and cognitive abilities of children with ASD (Mayo et al. 2013). Like age of first word, age of first sentence should be expected to be an important indicator of outcomes. The threshold for simple sentence (2-3 words) production in TD population is around 24-30 months. It has been suggested that in HF children with ASD (mean age of 9 years) the appearance of first sentence < 24 months as opposed to <36 months distinguished between those with better or worse adaptive communication (Kenworthy et al., 2012). Age of first sentence attainment has also been shown to be positively associated with nonverbal IQ (get from WISC-IV) at age 4 (Wodka et al., 2013). However several methodological limitations impact the conclusions that can be drawn from previous research. With the exception of the study by Kover and colleagues, studies have excluded children that experienced language regression/loss (Mayo et al., 2013),[8] and children with IQs in the intellectual disability range (i.e., < 70; Kenworthy et al., 2012), limiting the generalizability of the findings, particularly given the heterogeneity associated with ASD. Moreover, with the exception of Eigsti & Bennetto (2009), studies that reported on language outcomes as associated with both onset of single word and phrase speech, evaluated communicative and pragmatic abilities in children with ASD and did not investigate structural language

[8] Language regression means that following a period of typical language development, children enter a period in which already acquired language skills are no longer used (and may be lost) and acquisition of new skills declines (Barger et al., 2013; Kalb et al., 2010).

outcomes.

In the present work we aimed to address this lack of knowledge in the literature, comparing structural language abilities and nonverbal abilities of children with ASD, with both age of first word and age of first sentence, including children from the whole spectrum.

Research questions

Based on the literature reviews and methodological considerations presented in chapters 1 to 3, we asked the following research questions

1) **What are the measures most likely to appropriatley assess structural language and cogntive abilities in ASD?**

Our hypothesis in a nutshell (structural language abilities)

In Chapter 1 we hypothisized that repetition tasks, NWR and SR, and particularly linguistically based repetition tasks targeting computational complexity, both for phonology and morphosyntax, such as the LITMUS-NWR and the LITMUS-SR tests, should be the most reliable tools for evaluating structural language abilities in children with ASD, because these tasks narrow the possible influence of other factors (pragmatics, lexical knowledge, use of language in a conversational context, etc.) and allow for detection of children with structural language impairment both on a quantitative (general performance) and a qualitative (error analysis) level.

Background

The evaluation of language abilities in verbal children with ASD has typically been determined on the basis of vocabulary testing, while formal aspects of language in ASD (phonology and morphosyntax) have received much less attention. However, some studies have evoked the possibility that lexical ability involving knowledge of individual words can be a spared, or even enhanced, domain in children with autism (Walenski 2006 ; Mottron 2004). Moreover, some researchers have raised the possibility that neither standardized tests nor many experimental tasks targeting specific aspects of structural language may be sufficient to isolate the source of impairment in children with ASD (Kenworthy et al., 2012). In particular, measures drawn from standardized tests rarely enable detailed analysis of specific phonological and morphosyntactic structures. Recent studies have pointed out that low performance of children with autism on language tests targeting morphosyntax may be the result of a misunderstanding of the pragmatics of the testing situation, while tasks of articulation and repetition of real words may engage the use of a pre-existing lexical knowledge, which in relation to word familiarity or frequency effect can create biased performance. Some studies have also hypothesized that WM may play a large role in SR and NWR tasks. Mean Length of utterances (MLU) of sentences in SR and syllable length in

NWR can create memory effects on performance. However no study has tried to minimize possible effects of WM on both SR and NWR. Taken together, our discussion of the evaluation of morphosyntactic and phonological abilities pointed to the necessity of developing tools which both enable detailed analysis of formal language abilities that do not confuse linguistic performance with difficulties meeting task demands or with pragmatic impairment and with working memory effects, for morphosyntax, and narrow the possible influence of lexical knowledge, word frequency and syllable length for phonology. Moreover, the few studies which have used repetition tasks in children with autism distinguished two subgroups of language abilities (ASD-LI and ASD-LN) both on SR (Botting & Conti-Ramsden, 2003; Harper-Hill et al., 2013; Riches et al., 2010) and NWR (Bishop et al. 2004; Hill et al. 2015; Whitehouse et al. 2008, a.o.). We hypothesize that since these tasks are very sensitive to language impairment, they should also be able to detect structural language deficits among children with ASD. Pragmatic impairment typical of children with ASD should have limited impact on repetition accuracy in SR tasks (Polišenská et al. 2015; Silleresi et al. in press). The design of SR incorporates formal aspects of morphosyntactic processing but is less constrained by pragmatic features. Analogously, the design of a NWR task should ensure a more controlled evaluation of phonological abilities, since individuals cannot rely on their pre-existing lexical knowledge when repeating nonwords conceived to be unrelated to existing words in the child's language. We will seek controlled tasks of SR and NWR that can limit the effect of working memory and concentrate on syntactic and phonological constructions.

To our knowledge (except for one study, Sukenik, 2017, which looked at these phenomena indirectly), nobody has specifically investigated either the possible existence of enhanced lexical abilities in autism in relation to structural language abilities, or the possible differences in performance between standardized tasks evaluating formal language abilities and repetition tasks. We hypothesize that lexical abilities (assessed via tasks of vocabulary) may overestimate linguistic abilities in children with ASD and that standardized tasks evaluating phonological and morphosyntactic abilities may underestimate linguistic abilities of children with ASD. We predict that LITMUS-NWR and LITMUS-SR should be the best tools for evaluating structural language capacities in children with autism.

In order to verify our hypotheses we will seek answers to the following questions:

a. Do children with ASD display enhanced lexical abilities in relation to structural language abilities?

b. To what extent does the performance of children with ASD on standardized tasks evaluating phonology depend on previous lexical knowledge? To what extent does their performance on standardized tasks evaluating morphosyntax depend on difficulties meeting task demands or on pragmatic difficulties?
c. Do LITMUS-NWR and LITMUS-SR provide information on formal language abilities not provided by standardized tests?

Our hypothesis in a nutshell (cognitive abilities)

In Chapter 2 we assumed that NV cogntive tasks, RPM and the PRI of WISC-IV, can be the most reliable tools for evaluating cognitive abilities in children with ASD. These tools can be easily used in research and clinical practice; they highlight a full picture of autistic intelligence (strengths and weaknesses), and they can be used for an evaluation through the entire spectrum. Moreover, and this is important for the purpose of the present study, being NV tasks, they should not (or at least very remotely) be linked to verbal abilities. This should ensure reliable evaluation of the relation between linguistic abilities and cognitive abilities.

Background

There is evidence that the evaluation of cognitive abilities in children with ASD is hard to interpret. Some studies have found that the two main instruments used to evaluate cognitive abilities in individuals with ASD in the clinical domain, notably RPM in research and Wechsler scales (WISC), provide different results. While in typical populations there is homogeneity across both measures, in children with autism RPM scores have been reported to be significantly higher than FSIQ scores of the WISC. This is mainly due to the structure and the nature of the two tests. While RPM is a single test evaluating fluid and general intelligence, the Wechsler FSIQ score is derived from ten core subtests included in four different indices. Well-replicated findings suggest that nonverbal reasoning is a relative and perhaps absolute strength of autistic individuals. Because NV tests typically rely on abstract and fluid reasoning, children with ASD can leverage on their enhanced visuo-spatial abilities, without leaning on their diminished language or motor abilities. Discrepant performance between RPM and the Wechsler FSIQ has been demonstrated by recent studies reporting significantly higher RPM scores than FSIQ scores, in both high and low functioning children with ASD. Concerning the internal structure of the WISC, some studies found that the PRI of WISC-IV highlights the enhanced spatial perception and abstract reasoning abilities of individuals with ASD in comparison to both FSIQ and the other indices of the Wechsler scale.

However, even though RPM has been indicated as the most suitable test for evaluating autistic intelligence in research, and the PRI has been pinpointed as the most appropriate index of the WISC-IV for picturing autistic intelligence in the clinical domain, no study has compared these two measures. We hypothesize that since these tasks have been shown to enhance specific and different aspects of NV abilities, they can all be useful in a complementary evaluation of NV abilities of children with ASD.

Moreover, to our knowledge no study has directly verified the actual relation between linguistic abilities and performance on both RPM and the subtests of the PRI. Since in the present work we are arguing for the use of tasks that are nonverbal in nature (minimal or even absent need for task instructions or speech skills), we will verify the possible relations between RPM and linguistic abilities, and between each subtest of the PRI and linguistic abilities in children with ASD. Previous studies argued that RPM and two subtests of the PRI, Block Design and Matrix Reasoning, are not related to language (Houskeeper, 2011). We hypothesize that the Picture Concepts subtask, on the other hand, due to its nature, will be strongly related to language abilities. This hypothesis seems to go in the same direction as the new WISC-V version, which has excluded the Picture Concepts task from both the *Fluid Reasoning Index (FRI)* of the new primary index scales and from the new *Nonverbal Index* (NVI) of the ancillary index scales.

In order to verify our hypotheses we will seek answers to the following questions:

d. Do children with ASD display better performance on RPM and the PRI than on the FSIQ and all the other indices of the WISC-IV?

e. Do RPM and the PRI subtests evaluate different NV abilities or are they strongly correlated?

f. To what extent are RPM and the subtests of the PRI (Block Design, Matrix Reasoning, Picture Concepts) related to language abilities?

2) **When we apply the measures derived from the answer to (1) to a population sample taken from the entire spectrum, do clear language abd intellectual profiles emerge?**

Our hypothesis in a nutshell

We hypothesize that an integrative approach that takes in consideration both linguistic and cognitive abilities assessed through specific measures of formal language and NV cognition and that takes into account possible influence of extra-linguistic and extra-cognitive factors (severity of autism symptoms) may allow structural language/NV ability profiles to emerge from the totality of the (verbal) autism spectrum. From this analysis we expect that at least the four profiles of language/cognitive abilities detected in the literature will emerge in our group of children with ASD (ASD-LN with normal NVIQ; ASD-LN with low NVIQ; ASD-LI with normal NVIQ; ASD-LI with low NVIQ).

Background

To date, few studies have taken up the question of the nature of linguistic/cognitive profiles in children with ASD by explicitly exploring the interaction between language (dis)ability and intellectual (dis)ability and therefore their logically possible combinations. Typically, researchers have pointed to three profiles: two “homogeneous” profiles, ASD-LN with normal IQ and ASD-LI with low IQ, and one “discrepant” profile, ASD-LI with normal IQ, similar to the profile that defines children SLI (Geurts and Embrechts 2008). The existence of another “discrepant” profile, the ASD-LN with low IQ, has also been evoked (Joseph et al. 2002; Kjelgaard and Tager-Flusberg 2001; Tuller et al., 2017). Nevertheless, the relative infrequency of studies that have included low cognitive profiles in their population samples has resulted in a lack of knowledge about the capabilities of these children, and especially of the ASD-LN with low IQ profile (8% of participants in Joseph at al. 2002 and 10% of participants in Kjelgaard and Tager-Flusberg 2001). Since heterogeneity is a hallmark of ASD, a real understanding of linguistic/cognitive profiles can only be achieved through investigation of the entire spectrum, which is not restricted to Autistic Disorder (DSM-IV) and to high functioning autism.

Furthermore, few studies have investigated other factors that may be related to formal language or cognitive impairment in ASD. Notably, it remains to be seen if generalized measures of severity of autism symptoms might be related to poor performance of children in this population on formal aspects of language and/or intellectual disability. Current results have not yet given rise to a complete picture in the literature: some studies have found significant correlations between severity of autism and both language abilities and cognitive abilities (Gotham et al. 2009; Luyster et al. 2008), while others have found no such significant correlations (Charman et al. 2011; Loucas et al., 2008).

Applying measures of structural language abilities and NV intelligence derived from the answer to the first research question, we expect that the four profiles of structural language and NV cognitive abilities will emerge from an investigation extended the entire autism spectrum.

In order to verify our hypotheses we will seek answers for the following questions:

a. To what degree could extra-linguistic and extra-cognitive factors, e.g. severity of autism symptoms, and developmental factors, e.g. age of first word and age of first sentence, have an effect on structural language abilities and NV intelligence?
b. Which profiles of structural language/NV abilities can be detected in children with ASD?

3) Do the profiles identified in question (2) resemble the profiles of children with SLI and TD children?

Our hypothesis in a nutshell

We hypothesise that examining complexity of syntactic and phonological computation in ASD via the use of SR and NWR (tools specifically constructed to evaluate computational complexity), we could determine whether children with ASD-LI behave analogously to children with SLI, in avoiding structures that entail more complex derivations and in making errors related to complex constructions both on morphosyntax and phonology. We assume that if children with ASD-LI have the same phenotypical profile of children with SLI, they should present the same shortcomings in a quantitative, qualitative (error analysis) and developmental perspective. Concerning the ASD-LN profile we suppose that if structural language abilities of these children are really “normal” they should show a linguistic development similar to their typically developing age-peers.

Background

Although both ASD-LI and ASD-LN profiles have been detected in children with autism, the nature of linguistic profiles in ASD remains unclear. While some studies suggested that the phenotypical realization of linguistic shortcomings in ASD-LI is actually similar to the one found in SLI, qualitative error analyses showed that the two groups may behave differently on several constructions. Moreover, a few studies have recently showed that not all children with ASD-LN show a profile similar to TD age peers, suggesting that children with ASD-LN, although they have much less difficulty producing complex constructions than the two language-impaired groups, they may also show some (minor?) shortcomings. Some researchers have suggested that the interference of other factors (notably working memory effects, use of language in a conversational context, etc.) could obscur fundamental underlying similarities between the populations.

We propose an in-depth analysis of computational complexity effects, comparing the ASD-LI profile with children with SLI and the ASD-LN profile with TD children in order to cast new light on the debate on similarities versus differences between these groups. Since the tools we used for the evaluation of morpshosynatctic and phonological abilities have been identified as the most suited for evaluating children with ASD and they were built for minimizing possible effects of factors other than structural language abilities, and since profiles of structural language / NV abilities obtained as a result of question (2) were controlled for internal homogeneity, the comparison with SLI and TD groups should be "cleaner" than such comparisons typically are.

In order to verify our hypotheses we will seek answers for the following questions:

a. How similar are the linguistic capacities of ASD-LI to those of children with SLI on phonology and on morphosyntax?
b. How spared are the language abilities in children with ASD-LN on phonology and morphosyntax?

To answer both these questions we will examine performance of children with ASD-LI and ASD-LN in comparison to children with SLI and two group of younger and age-matched TD children, via a quantitative (global performance), qualitative (error analysis) and developmental analysis of computational complexity effects on SR and NWR.

Part II

Experimental study

Chapter IV
Methodology

4.1 General methods

The aim of the present study was to put the heterogeneity of ASD to the forefront by investigating whether clear profiles related to structural language and NVIQ abilities would emerged when investigation is extended to the entire spectrum. In Part II of our work we will limit our analysis to monolingual children with ASD. In the next section we will report the main characteristics of the included populations. We will come back on bilingual children in Part III.

4.2 Inclusionary criteria and recruitment procedures

Participants with ASD were recruited from the Autism Centre at the Regional University Hospital Centre in Tours (France). To participate in the study, all participants had to meet the following criteria:

(1) A prior DSM-5 clinical diagnosis of ASD by the child's health care professionals. The diagnostic classification was based on diagnostic assessment by a child psychiatrist and multiple informants (i.e., speech-language pathologists, psychologists, and educators) and confirmed, when possible, by the ADOS module 1(few words), 2 or 3, and the ADI-R. All children presenting other co-occurring conditions, especially medical or genetic conditions and other neurodevelopmental disorders, were excluded from the present study (no neurological, sensory, or motor impairment; no hearing impairment; no extreme language deprivation or unfavourable language environment).

(2) All children in the present study were monolingual French-speaking children.

(3) We chose to include children between 6 and 12 years of age for three main reasons. First, past studies Bennett et al. (2014) showed that for children with ASD, language impairment measured at ages 6–8 years old was a more robust predictor of symptoms and functional outcomes than language measured at earlier ages (2-4 years old) (Bennett et al., 2008). Second, since some studies demonstrated improvement with age of linguistic abilities in children with ASD (Anderson et al., 2007; Geurts & Embrechts, 2008) we decided to include children with a wide age range, in order to

verify the possible relation of linguistic abilities and developmental trajectories. Third, children at age 6 would already have acquired the vast majority of simple and complex syntactic constructions making it possible to test more aspects of syntactic functioning.

(4) A minimum MLU (mean length of utterance) of 2.5, an inclusionary criterion to ensure that language tests could be administered.

(5) No inclusionary criterion was set for IQ, as one of our goals was to look at the nature of formal language in verbal children across the spectrum, and, this measure indeed varied.

All children meeting the previous criteria and present at the hospital centre during the recruitment phase of the study were offered the opportunity to participate. A formal letter explaining the research aims and protocol was sent to the parents of all the children; parents had to sign a consent form enabling their child to participate in the research. Children were recruited through two main channels inside the Autism Centre at the Regional University Hospital Centre in Tours: the day-care section (Day Hospital) and the autism diagnostic section (Autism Resources Centre – CRA). Although 68 monolingual children with ASD could have potentially participated to our study (letters were sent to each child's parents), data were gathered for 38 children (24 children were recruited from the day-care section and 14 from the CRA). Thirty children were excluded from the final clinical group for different reasons: 12 parents did not return their consents; for 8 children it was impossible to administer the protocol because they were too tired, too distracted or too severely impaired (not enough language to perform the entire protocol); 4 children did not come to their appointment for receiving diagnosis at the CRA; 3 children received a diagnosis other than ASD at the CRA; 2 children were transferred to other centres; and 1 child was already included in another research protocol. The different organization of the two sections of the Autism Centre in Tours influenced both the mean of recruitment and the number of exclusions from our protocol. Children at the Day Hospital section, who were already diagnosed with ASD, were hospitalized and divided into intervention groups, depending on age, frequency of intervention (which could range from four days per week – to one session every two weeks) and therapeutic needs (speech-language therapy, psychomotor-therapy, psychotherapy, a.o.). The presence of these children at the hospital facilitated our recruitment and supplied the possibility of modelling, based on the specific needs of each child, the number of sessions for the administration of our protocol. This led to the relatively low rate of exclusion of children from the day-care section (25% of the total sample recruited from the day-care section). On the other hand, children recruited from the CRA section came to the centre exclusively to receive a positive or negative diagnosis of ASD and they were seen only once (for one day only). Our research protocol was systematically administrated, in this case, at the end of the day, after the series of clinical exams necessary for the diagnosis. Children were often tired, distracted or simply too severely impaired to perform our protocol. In this case no other complementary session could be administrated, which led to many children being excluded (75% of the total sample recruited from the CRA section).

The assessment battery used in the study was administrated by the same researcher (the author of the present work) or by one student/research assistant. Since the assessment criteria of the day hospital and the CRA sections were the same, assessment consistency was ensured across participants.

4.3 Ethics statement

The study was carried out with the approval of the Ethical Committee for Non-Interventional Research (CERNI) of Tours-Poitiers (France). The research plan was approved by the head of the Autism Centre at the Regional University Hospital Centre in Tours, Pr. Bonnet-Brilhault, by the University of Tours and by the doctoral school. To preserve the anonymity of the patients, all children were assigned a randomly self-generated three letter code, with which they will be identified throughout.

4.4 Participant characteristics: clinical data

4.4.1 Age, sex and diagnosis

Thirty-seven verbal children with ASD, aged 6-12 years old (*M* = 8;8, SD = 18.9, range = 6;2 -12), were recruited from the Autism Centre at the Regional University Hospital Centre in Tours (France). All participants were monolingual French-speaking children and met the criteria for a DSM-5 clinical diagnosis of ASD, confirmed by the ADOS module 1 (few words), 2 or 3, and/or the ADI-R. For the purpose of present study (evaluating language abilities in children with ASD) we excluded children that were assessed with ADOS Module 1 (no words) because these children were nonverbal.

Clinical information about participants, including diagnosis and autism severity scores, were collected from the hospital clinical database. Using the ICD-10 criteria retrieved from the database, the group was composed of 15 children with Autistic disorder (A), 13 children with PDD-NOS (P) and 9 children with Asperger's Syndrome (AS).[9] The group was composed of 1 girl and 36 boys. Table 6 presents the main characteristics of the total sample of participants. Empty cells indicate information that was missing from the clinical database. Data from the ADI-R were missing for seven children. However, since these children had a diagnosis of autism via the ADOS global score, confirmed by psychiatrists at the Autism Centre in Tours, they were included in the present study.

[9] The latest draft of the manual, the ICD-11 published in May 2018 by the World Health Organisation (WHO), collapses the three previously distinct diagnoses of Autism, Asperger and Pervasive Developmental Disorder-Not Otherwise Specified (PDD-NOS) into a single diagnosis of 'Autism Spectrum Disorder.' This change mirrors the criteria of the fifth edition of the "Diagnostic and Statistical Manual of Mental Disorders" (DSM-5) released in 2013. However, since all the participants in the present study were recruited before the release date of the ICD-11 manual, we think it is important to specify the classification assigned by the ICD-10, then applicable.

TABLE 6. ASD PARTICIPANTS' CHARACTERISTICS

	Child code	Gender	Age (y;m)	Clinical Diagnosis (ICD-10)	ADOS module
1	SIM	M	6;3	Autism	1
2	SCO	M	6;8	PDD-NOS	2
3	RUG	M	6;9	Autism	1
4	ASC	M	6;11	Autism	2
5	KEV	M	7;0	PDD-NOS	2
6	VOR	M	7;2	Asperger	2
7	MTH	M	7;3	PDD-NOS	1
8	YLA	M	7;4	PDD-NOS	2
9	LWA	M	7;5	Autism	1
10	ODI	F	7;5	PDD-NOS	1
11	EVA	M	7;8	PDD-NOS	1
12	MON	M	7;9	Autism	2
13	CUT	M	8;0	Asperger	2
14	CIP	M	8;1	PDD-NOS	2
15	NUG	M	8;3	Autism	1
16	ROS	M	8;3	Autism	1
17	NAF	M	8;4	Autism	2
18	NOS	M	8;4	Asperger	2
19	ATE	M	8;9	PDD-NOS	3
20	EDT	M	8;10	Autism	3
21	EPI	M	9;2	PDD-NOS	1
22	EMP	M	9;2	PDD-NOS	3
23	TUC	M	9;3	Asperger	3
24	GOT	M	9;5	Asperger	3
25	ELO	M	9;6	Autism	2
26	AVI	M	9;9	Asperger	3
27	MUG	M	9;9	Autism	3
28	JOS	M	9;10	Autism	1
29	FIZ	M	9;11	PDD-NOS	2
30	GHO	M	9;11	Asperger	3
31	ADO	M	10;9	Autism	2
32	LPG	M	10;10	Asperger	3
33	MIR	M	11;3	Autism	1
34	LAT	M	11;6	Autism	3
35	MOI	M	11;6	PDD-NOS	3
36	YAT	M	11;8	PDD-NOS	1
37	LEC	M	12;0	Asperger	2

4.4.2 *Severity of autism scores*

Following the criteria that we introduced in Chapter 3, we gathered data for severity of autism scores from the ADOS calibrated severity score (standardized by Gotham et al., (2009) from the revised algorithms of the raw total scores, the CARS global score, and the ECA-R global score. The cut-off for displaying typical autistic traits is ≥ 4 for the ADOS and ≥ 30 for the CARS. The ECA-R does not provide any cut-off score. The CARS and ECA-R scales were also used, in the present study, to retrieve specific subscores of autism severity in relation to particular areas of development (e.g. the score for echolalia that will be used in analysis of results in Chapter 5).

Table 7 presents the autism severity scores for the 37 children that met the previous criteria for a diagnosis of ASD. Since autistic symptomatology is known to evolve over time, we looked at the most recent evaluations run before the administration of our protocol for all three autism severity scores. For the ADOS, this score often corresponded to the date of the diagnosis; for the CARS and the ECA-R, the gap between the administration of our protocol and these scales was less than twelve months.

TABLE 7. AUTISM SEVERITY SCORES OF CHILDREN WITH ASD (ADOS SEVERITY SCORES, CARS TOTAL SCORES AND ECA-R TOTAL SCORES)

	Child code	Gender	Age (y;m)	Clinical Diagnosis (ICD-10)	ADOS severity score	CARS	ECA-R
1	SIM	M	6;3	Autism	6	21	27
2	SCO	M	6;8	PDD-NOS	6	38.5	31
3	RUG	M	6;9	Autism	6	34	10
4	ASC	M	6;11	Autism	6	23	27
5	KEV	M	7;0	PDD-NOS	9	29.5	27
6	VOR	M	7;2	Asperger	2	22	17
7	MTH	M	7;3	PDD-NOS	8	30	34
8	YLA	M	7;4	PDD-NOS	2	22	14
9	LWA	M	7;5	Autism	9	30	38
10	ODI	F	7;5	PDD-NOS	8	26	5
11	EVA	M	7;8	PDD-NOS	6	28	35
12	MON	M	7;9	Autism	4	24	45
13	CUT	M	8;0	Asperger	8	28	16
14	CIP	M	8;1	PDD-NOS	4	24	22
15	NUG	M	8;3	Autism	7	33	11
16	ROS	M	8;3	Autism	4	26	13
17	NAF	M	8;4	Autism	8	31	42

18	NOS	M	8;4	Asperger	8	31.5	34
19	ATE	M	8;9	PDD-NOS	4	27	24
20	EDT	M	8;10	Autism	2	26	15
21	EPI	M	9;2	PDD-NOS	6	28	9
22	EMP	M	9;2	PDD-NOS	6	25	5
23	TUC	M	9;3	Asperger	7	22	13
24	GOT	M	9;5	Asperger	3	23	16
25	ELO	M	9;6	Autism	2	27	14
26	AVI	M	9;9	Asperger	5	19.5	12
27	MUG	M	9;9	Autism	4	26.5	9
28	JOS	M	9;10	Autism	6	32.5	34
29	FIZ	M	9;11	PDD-NOS	7	33.5	42
30	GHO	M	9;11	Asperger	8	30	34
31	ADO	M	10;9	Autism	7	29.5	68
32	LPG	M	10;10	Asperger	2	30	8
33	MIR	M	11;3	Autism	4	25	21
34	LAT	M	11;6	Autism	10	27	23
35	MOI	M	11;6	PDD-NOS	4	21.5	14
36	YAT	M	11;8	PDD-NOS	5	31.5	38
37	LEC	M	12;0	Asperger	6	20	2

Note: cut-off for typical autistic traits is ≥ 4 for the ADOS and ≥ 30 for the CARS

Looking at the severity scores of the 37 children with ASD included in our study, we can see that for 6 children (VOR, YLA, EDT, GOT, ELO, and LPG) the ADOS scores indicated the presence of very few autistic traits at the time of assessment (a severity score < 4). More dramatically, on the CARS scores, 25 children did not display typical autistic symptomatology (a severity score < 30) at the time of our assessment. However, not showing typical autistic symptomatology does not mean that these children grew out of the spectrum. This feeble presence of symptoms may be the result of early intervention programs as demonstrated by Camaioni and colleagues (2003) and Rogers (1996), who assessed autism severity via CARS scores in a longitudinal perspective. We will carefully consider the CARS scores in our analyses.

4.4.3 Developmental factors

Information about age of first word and age of first sentence (both in months) were collected from the clinical database. However while information about age of first word (*M* = 22.5, SD = 9.2, range = 8 - 48) was available for almost the totality of the participants (34/37 children), age of first sentence (*M* = 39.8, SD = 17.6, range = 15 -84) was retrievable for only 23 children. Measures with missing data points will be used with caution in further analyses.

4.5 Research protocol

Children were evaluated on language abilities via a battery of standardized language tasks including vocabulary, morphosyntax and phonology, via experimental repetition tasks targeting specific aspects of structural language (SR and NWR), and on cognitive abilities via both RPM and a standardized battery of cognitive abilities (when possible, data from the WISC-IV battery were retrieved from clinical records).

4.5.1 Language Measures

4.5.1.1 Standardized language tasks

Children in the ASD group were assessed in three areas of language ability: vocabulary, phonology and morphosyntax, with tools commonly used by speech-language pathologists (SLP) in France. Children recruited from the day-care section of the Autism Centre were evaluated on receptive vocabulary via a classical picture-pointing task, taken from the Evaluation de Langage Oral, ELO, battery (Khomsi, 2001). Children recruited from the CRA, were assessed on receptive vocabulary via the EVIP, the French adaptation of the Peabody Picture Vocabulary Test (Dunn et al., 1997) This difference was due to the fact that in the CRA section our evaluation of language abilities through the use of standardized tests was integrated into the clinical assessment of the SLP in charge of the children's diagnosis, who preferred to use the EVIP. Expressive vocabulary was assessed via the corresponding subtest of the ELO or of the BILO, Bilan Informatisé de Langage Oral (Khomsi et al. 2007), the computerized version of the ELO battery. In this task children are asked to name two series of pictures, one representing concrete and abstract objects and the other representing a series of verbs. Phonology was evaluated via the word repetition task of the BILO, which consists of 42 pre-recorded words of increasing length and phonological complexity, some of which are unfamiliar to children, and thus serve as quasi-nonwords, according to the authors of the test. Both Receptive and Expressive morphosyntax were evaluated via subtests of the BILO. Comprehension test consisted in a 22 items picture-matching task, while sentence completion test, entailed 25 items in which children had to complete a sentence on the basis of a picture and a verbal stimulus. These tasks will be described in detail in Chapter 5 (section 5.2.3.2).

For all standardized tests, raw scores were converted into z-scores in order to take into account population norms. Following Tomblin (2011), who motivates the cut-offs with an explicit comparison to a gold standard of experienced SLP assessment of spontaneous language samples, we established a cut-off for moderate impaired performance at -1.25 SD (the mean rate of Tomblin's suggestion) for all standardized tests. However in clinical setting in France the cut-off of -1.65 SD (5th percentile) is often used to indicate severe language impairment (Ramus, 2003). For these reasons we tried to remain as neutral as possible choosing to adopt both cut-offs for describing the severity of language impairment in the performance of children with ASD on standardized tests. We will use -1.25 SD for determining language impairment, and we will indicate performance between -1.25 and -1.65 SD as moderate language impairment and performance < -1.65 as severe language impairment.

4.5.1.2 Experimental language tasks

As anticipated in Chapter 1, in our study we argue for the use of two linguistically based repetition tasks evaluating phonology and morphosyntax, the LITMUS-NWR-French (Ferré & dos Santos 2015) and the LITMUS-SR-French (Prévost, Tuller & Zebib 2012). These tools were developed during COST Action IS0804 (European Cooperation in Science and Technology), devoted to bilingualism and SLI. The action's aim was to create possible diagnostic tools and/or screening tools able to disentangle typically developing monolingual and bilingual children from monolingual and bilingual children with SLI. Both these tasks evaluating specific aspects of phonology and morphosyntax demonstrated to be very sensitive to SLI (Conti-Ramsden et al., 2001). In the present work we wanted to apply these tools to the ASD population in order to see if they are able to disentangle children with ASD and normal language abilities and children with ASD and language impairment. In this section we will describe both tasks, highlighting how they answer the limitations displayed by other repetition tasks used in previous studies (see sections 1.2.10.2 and 1.2.11.2). Both LITMUS tasks are presented in Appendix 1.

LITMUS-NWR-French

LITMUS-NWR-French (Ferré & dos Santos, 2015) elicits the repetition of 50 nonwords. This task was designed to focus on syllable complexity (e.g., the presence of different types of consonant clusters and word-final consonants), and the stimuli were controlled for length and word-likeness. The following description of the tasks was taken from dos Santos & Ferré, (2018). The nonword repetition test was built to test only phonology as much as possible. The main aim of the test was to limit two external influences often encountered in this type of tasks, **working memory** and **lexical knowledge**, which have been identified as the most important limitations of previous NWR tasks (Chapter 1). For **working memory**, the longer the nonword, the stronger is the effect of working memory on repetition performance. The effect of length starts to appear in children with SLI when nonwords have more than two syllables (section 1.2.10.2). In most NWR tasks, nonwords can have up to five syllables and sometimes even more (e.g., Archibald & Gathercole, 2007; Bishop, 2004; Riches et al., 2011). Since one of the aims of the present study was to compare children with ASD with children with SLI in order to see if there was a subgroup of children with ASD whose language ability profile resembled that displayed in SLI, we needed to limit the effect of working memory. In this way the performance of children with SLI should be less dependent on nonword length and the comparison with ASD should therefore be more reliable. For these reasons we chose to use LITMUS-NWR. In this task nonwords have no more than three syllables. **Lexical knowledge** can also play a role when nonwords sound like real words. This fact can be captured with a measure of word-likeness (Frisch et al., 2000; Munson et al., 2005). In order to limit the effect of lexical knowledge, nonwords were built through the use of elementary units (syllables and segments), combined to create nonwords. In this way the blocks could be manipulated on syllable structure, which seemed to be a suitable marker for assessing phonological disorders (see Ferré et al., 2012 for a detailed description).

The test was constructed to assess specific aspects of phonological and to assess phonological skills in different languages since it was created for disentangling monolingual and bilingual children with SLI from TD population. In order to cover these prerequisites, two types of items were created: the so-called “language independent” items (LI, n = 21) and “language dependent” items (LD, n = 29). The LI items had a phonological structure that was possible in most languages of the world. This label was given for the sake of simplicity. These items should be more properly characterized as “quasi-independent.” Indeed, it is not possible to completely isolate nonwords from lexicon (Chiat, 2015). To create LI items, three aspects of phonological complexity were selected.

The first one was syllable structure. Three different syllable structures were involved. The simplest syllable structure consisted of a consonant followed by a vocalic element: CV.

This syllable type is universal since it is present in all languages examined to date (Goad & Rose, 2004). Two other syllable types were selected as well (CCV and CVC#) because of their complexity compared to CV syllable and because these syllable types are found in a majority of the world's languages. According to Maddieson, (2009) in a sample of 515 languages, 88% can have a branching onset and/or consonant after a vowel, generally at the end of a word.

The second aspect was related to segmental complexity. Segmental complexity was taken into account for consonants only. The consonants that were selected are early acquired and available in a large number of phonological systems. There were two stops, [p] and [k], one fricative [f] and one liquid [l]. As a dorsal consonant, [k] is considered more complex than [p], which is labial (Hayes, 2004). As for fricatives, they are considered more complex than stops (Jakobson 1969). Finally, [l] was selected in order to have items with branching onsets, which are present in a large number of languages. For vowels, the three most common vowels among the languages of the world were selected: [i], [a], and [u] (Maddieson, 2009).

The third aspect was sequentiality (i.e., the syntagmatic axis). Two types of sequentiality were hypothesised to increase item complexity: consonant sequence and syllable sequence. As regards consonant sequences (consonant clusters), when places and/or manner of articulation alternate ("pu**sk"**), the sequence was viewed as more complex. Furthermore, items beginning with a labial were regarded as simpler ("**p**ufaki"). For syllable sequences, in three-syllable items, the second syllable was more fragile in French, and even more so when this syllable involves one or more aspects of complexity (e.g., a branching onset "ku**fl**api"). In other languages, when stress is controlled, it seems to be the case as well. For example, Marshall & Van der Lely, (2009) found that children with SLI and children with dyslexia produced more errors on word-medial clusters compared to word-initial clusters.

LD items were created by integrating two aspects of complexity that are part of French phonology (and also present in other languages). LD nonwords were created with the same syllable types and segments as in LI nonwords, to which the coronal fricative [s] was added. This additional consonant enabled the creation of nonwords with the following complex consonant clusters: #sCV, #sCCV, sC#, and Cs#. These consonant clusters are unusual in the world's languages (Goad & Rose, 2004) but are allowed in French. For the creation of LD items, internal codas were also included ("paklu"), which was not the case for LI items.

In both the LI or LD items, the frequency of occurrence of each sound was controlled for. Indeed, [p] is as frequent as [k], and [i] is as frequent as [u]. The authors also controlled for wordlikeness, comparing the items to each other. They found that none of the items was more wordlike than the others. All nonwords (LI and LD) were randomly ordered. The last step was to reorder consecutive nonwords that were too close phonetically. Illustrative items for each condition present in the task are reported in Table 8.

TABLE 8. THE LITMUS-NWR-FRENCH

Control items			
CCV	[plu]	CVC	[kip]
sCV	[spu]		
Low complexity items			
CV.CCV	[paklu]	CCV.CV	[plifu]
CV.CVC	[pukif]	CV.CVs	[fapus]
CV.CV.CV	[kifapu]		
Average complexity items			
CCV.CVC	[flukif]	CCV.CCV	[flaplu]
CCV.CV.CV	[flipuka]	CV.CV.CCV	[kupifla]
CV.CCV.CV	[kuflapi]	CV.CV.CVC	[kapufik]
CV.CV.CVs	[pifukas]		
CCVL	[plal]		
CVCs	[fips]	CVsC	[pusk]
CCVs	[flis	sCV.CV	[skafu]
sCCV	[skla]	sCVC	[skap]
CVL.CV	[pilfu]	CVs.CV	[kusp]
High complexity items			
CCVCs	[pliks]	CCVsC	[plusk]
CV.CVL.CV	[kufalpi]	CV.CVs.CV	[pafuski]
sCV.CV.CV	[skapufi]		

For **LITMUS-NWR the repetition cut-off rate was established at 77%,** which corresponded to high levels of diagnostic accuracy for language impairment (specificity, 83%, and sensitivity, 88%) in Tuller et al.'s (2018) study (obtained from a group of 5;6 to 8;6-year-old monolingual TD children and children with SLI). We predicted that computational complexity would have a stronger effect on children with ASD that display language impairment than in children who show normal linguistic abilities. For LITMUS-NWR we expected that low complexity items would be repeated better than more complex ones and that syllable length would not play a role in the performance of children with ASD.

LITMUS-SR–French

LITMUS-SR-French (Prevost, Zebib, & Tuller, 2012) was composed of 30 sentences which varied in syntactic complexity, including morphosyntactic properties known to be difficult for children with SLI: verbal morphology and complex syntax (with movement and/or clausal embedding). The sentence repetition test was built to assess, as much as possible, only morphosyntactic abilities. We chose to use this test in order to overcome the limitations found in any tasks identified in section 1.1.11.2. None of the previous studies utilized a linguistically based sentence repetition task that tested computational complexity including a variety of different structures and that limited the use of short-term memory.

Computational complexity in the generative linguistic theoretical framework can be measured in terms of the nature and number of operations needed for the derivation of a syntactic construction. Children with language impairment, notably children with SLI, show a deficit in the computational system, which leads to the inconsistent use of certain grammatical operations. It has been demonstrated that structure-dependent relationships that can be observed in tense marking, case marking and more importantly in long-distance dependencies which necessitate movement (wh-questions and relative clauses) and in embedded clauses (argument clauses and relative clauses) may constitute sources of impairment for children with SLI. If we look at previous studies on language in ASD, only Riches et al. (2010) employed an experimental task constructed to specifically evaluate computational complexity. However, while their SR task met the criterion of computational complexity, it focused only on relative clauses, which prevented investigation of a variety of complex constructions and excluded simple sentences altogether. Having constructions other than relatives in the SR task, computationally less, equally or more demanding could be useful for detecting potential difficulties and errors in groups with language impairment and therefore more appropriate for comparing children with ASD-LI and SLI on a variety of structures. Finally, a more heterogeneous task would be easily transferable to clinical practice since it should be less stressful for children, who should at least succeed in repeating less complex sentences, and it should be easier to use with a wide variety of children with ASD. The LITMUS-SR task was constructed in order to have both less complex (computationally less demanding) and more complex structures (computationally more demanding) structures. Specifically, the task contained five sentence structures each divided into two substructures, one less complex and the other more complex, consisting of three sentences each (for a total of 30 sentences). Two additional sentences were used for training.

There were two types of monoclausal sentences with canonical word order (SVO) varying in tense (present vs. past) and in number marking on the verb (singular vs. plural):
(i) Monoclausal sentences in the **present tense**, with either a third person singular **3S** (less complex) or a third person plural subject **3P** (more complex)
(ii) Monoclausal sentences in the **past tense**, with either a third person singular **3S** (less complex) or a third person plural subject **3P** (more complex)

There were three types of clauses involving movement and/or embedding: monoclausal object wh-questions with either a non-discourse-linked operator (*qui* 'who') or a D-linked operator (*quel* 'which'), biclausal sentences with either a nonfinite or finite complement clause, and biclausal sentences with either a subject or object relative clause.
(iii) **Object wh-questions**, who-questions **Qui** (less complex) and 'which N' questions **Quel** (more complex)
(iv) **Argument clauses**, with nonfinite verb **Nonfinite Arg. clauses** (less complex) and with finite verb **Finite Arg. clauses** (more complex)
(v) **Relative clauses**, subject relatives **SR** (less complex) and object relatives **OR** (more complex)

The distinction between less complex and more complex substructures was made on the basis of generative syntactic theory. Concerning monoclausal SVO sentences, difficulties with plural verbal agreement were reported in French-speaking TD children and in children with SLI by Franck et al., (2004). The authors suggested that plural subject-verb agreement involves some level of higher computational complexity than the singular subject-verb agreement. This complexity is due to markedness and extra-feature checking of 3rd person plural condition. Singular verbs are often produced by default and since plural agreement entails computation of feature checking (probe/goal Spec-head relationship), it is highly error-prone. The language systems of young TD children below 8 years old seems to show this singular-plural asymmetry. Children with SLI, on the other hand, were found not only to have failed to develop such an automatic computation of feature checking, but they also seem to show a deviance in elaborating structural dependencies (Franck et al., 2004).

Regarding the other types of sentences included in the task, the literature has widely demonstrated that both TD children (up to age 9) and children with SLI experience difficulties with A' movement, such as object which-questions, as opposed to object who-questions and object relative clauses as opposed to subject relative clauses (Friedmann et al., 2009; Novogrodsky et al., 2006; Van der Lely, 1996). Several studies have analysed the asymmetry between *which/who* object questions (Hamann, 2006; Jakubovitz & Tuller, 2008; a.o.) and OR/SR (Delage et al., 2008; Guasti & Cardinaletti, 2003) in French-speaking TD children and children with SLI. Difficulties in parsing the dependencies of both object *which*-questions and object relatives have been explained through the theoretical framework of Relativized Minimality (RM) (Rizzi, 1990). RM states that in the configuration (1), a local relation cannot hold between X and Y when Z intervenes. In particular, Z is considered a strong intervener when it shares a significant number of features with X.

(1) X … Y … Z

In object *which*-questions, as in object relatives (OR), a lexically restricted ([+LR]) subject plays the role of an intervener in a dependency relation between the moved [+LR] *which*-N object (the object DP in OR) and its trace, thus increasing computational complexity. Intervention does not play a role in object *who*-question since *who* is [-LR], and in subject relatives (SR) since there is no intervener between the gap (the initial position) and the head of the relative. These differences are illustrated in (2) for wh-questions and in (3) for relative clauses.

(2) Which- a. [Quel enfant]$_i$ la maitresse punit t_i ? More complex
[Which boy]$_i$ the teacher punishes t_i?
Wh NP D NP <Wh NP>

Who- b. [Qui]$_i$ la maitresse punit t_i ? Less complex
[Who]$_i$ the teacher punishes t_i?
Wh D NP <Wh>

(3) OR a. Tu as vu [le cheval]$_i$ que le chien a mordu t$_i$ More complex

You saw [the horse]$_i$ that the dog has bitten t$_i$

D NP R D NP <D NP>

SR b. J'ai vu [le chat]$_i$ qui t$_i$ a griffé la vache Less complex

I saw [the cat]$_i$ that t$_i$ has scratched the cow

D NP R <D NP> D NP

Finally finite complement clauses have been found to be particularly problematic, in comparison with nonfinite complement clauses, for both TD children and children with SLI (Hamann et al., 2007; a.o.). They involve increasing computational costs due to the presence of Complementizer-Tense agreement and subject-verb agreement, mood and tense dependencies between the matrix and the embedded clause, and overt subjects and overt complementizers which agree in Wh-features with the force of the clause (Jakubowicz & Tuller, 2008), as shown in (4)

(4) a. Le papa$_i$ sait très bien [$_{CP}$ Ø [$_{IP}$ PRO$_i$ conduire la voiture]].

the daddy knows very well drive-inf the car

b. La dame dit [$_{CP}$ que [$_{IP}$ le garçon a pris le ballon]].

the madam says that the boy has taken the ball

Concerning short-term memory, SR tasks are typically highly memory based. In the most frequently used SR task in published studies on ASD, the Recalling Sentences of the CELF, sentences can be up to 18 words (23 syllables) long, which makes it very difficult to tell whether low performance is due to the complexity of the structures or to demands on short-term memory. Since in the present study we aimed to focus on the errors that can stem from the complexity of syntactic structures in order to describe the profiles of structural language abilities in children with ASD, in comparison to children with SLI, the effect of working memory needed to be limited. In particular, performance should not be dependent on sentence length. The LITMUS-SR task was constructed to control for this particular aspect. In this task sentences have no more than 10 words (13 syllables). Moreover, within each structure condition (less complex and more complex substructures), the sentences did not differ significantly in terms of number of syllables. This means that any performance differences in the repetition of more vs. less complex structure subtypes should not be attributable to memory components, but rather should reflect the effect of computational complexity. The length of monoclausal sentences, including object wh-questions, ranged from 5 to 9 syllables, whereas for biclausals length ranged between 10 and 13 syllables. A presentation of the LITMUS-SR task is given in Table 9.

For LITMUS-SR the repetition cut-off rate was established at 78%, which corresponded to high levels of diagnostic accuracy for language impairment (specificity, 92%, and sensitivity, 93%) in study (obtained on monolingual TD children and children with SLI). We predicted that computational complexity would have a stronger effect on performance of children with ASD that display language impairment than in children who show normal linguistic abilities. We also expected that repetition accuracy should be higher for less complex substructures than for more complex ones.

TABLE 9. THE LITMUS-SR-FRENCH

Structure	Comparatively LESS complex structures		Comparatively MORE complex structures		Mean length (# syllables) and range
SVO present tense	3S	*Le garcon prend un bain* the.masc.sg boy takes a.masc bath 'The boy takes a bath'	3P	*Les chats boivent du lait* the.pl cats drink.3pl some.masc milk 'The cats drink some milk'	6.7 (1.0) 5–8
SVO past tense	3S	*La maman a fermé la fenêtre* the.fem mother has closed the.fem window 'The mother closed the window'	3P	*Les parents ont rangé les jouets.* The.pl parents have.3pl put+away the.pl toys 'The parents put away the toys'	8.7 (0.4) 8–9
Wh-object questions	Qui	*Qui la maîtresse punit?* Who the.fem.sg. teacher punishes? 'Who does the teacher punish?'	Quel	*Quel enfant la maîtresse punit?* Which child the.fem.sg. teacher punishes? 'Which child does the teacher punish?'	7.0 (1.0) 6–8
Argument clauses	Non-finite	*Le papa sait très bien conduire la voiture* The.masc.sg. daddy knows very well drive.inf the.fem.sg. car 'The daddy knows to drive the car very well'	Finite	*La dame dit que le garcon a pris le ballon* The.fem.sg. woman says that the.masc.sg. boy has taken the.masc.sg. ball 'The woman says that the boy has taken the ball'	11.8 (0.3) 11 – 12
Relative clauses	SR	*J'ai vu le chat qui a griffé la vache* I've seen the.masc.sg. cat who has scratched the.fem.sg. cow 'I've seen the cat who has scratched the cow'	OR	*Tu as vu le cheval que le chien a mordu* You've seen the.masc.sg. horse whom the.masc.sg. dog has bitten 'You've seen the horse whom the dog has bitten'	11.3 (0.7) 10 – 13

4.5.1.3 Cognitive tasks

Following the criteria introduced in Chapter 2, we gathered data about the cognitive level of the children with ASD from the clinical database. Out of the 37 children that were included in our protocol, 33 had been administrated the WISC-IV battery. We collected scores for the FSIQ, all four indices, and for the Block Design, Matrix Reasoning and Picture Concepts subtests. The gap between the administration of the WISC-IV and our protocol was less than twelve months (M = 2.72, SD = 8.9) for each participant, and the two protocols were systematically administered in that order. The four remaining children had scores from another cognitive evaluation, the Echelles Differentielles d'Efficience Intellectuelle, EDEI-R (Perron-Borelli, 1996). For these children we collected scores for FSIQ, Verbal efficacy and NV efficacy. Finally, additional information on the children's NV intelligence was gathered via RPM, which were included in the protocol. For the present study we chose the RPM version in the colored format with 36-items. For the description of all cognitive tasks used in the present work, see Chapter 2.

The cut-off for intellectual impairment on RPM was established at < 10^{th} percentile, which corresponds to < 80 standard score, following the criteria of the Ravens' manual (Raven, 1998). For comparative purposes we used the same cut-off for the WISC-IV (FSIQ, indices and subtests) and the EDEI-R. Moreover, the choice of 80 standard score as cut-off for intellectual impairment seems to go in the same direction as the new WISC-V version. In the descriptive classification of IQ levels, the most recent version of the Wechsler scale has changed the description of the scores' band between 70 and 79 from "borderline IQ" to "very low IQ", indicating this range of scores as impaired (Raiford & Holdnack, 2014).

4.6 SLI group and TD groups

The target ASD group was compared to a group of 26 chronologically age-matched children with SLI (age range: 6;2-11;1, M = 8;6, SD = 17.5) and to 84 younger and age-matched TD children. The children with SLI were recruited via the university teaching hospital centre specialized in language and learning disability diagnosis in Tours. They were diagnosed following usual exclusionary criteria, and thus, notably, they had normal nonverbal intelligence and autism had been ruled out by a psychiatric examination. At the time of diagnosis, all of these children were deemed to be impaired in both phonology and syntax, a constellation commonly referred to in the French clinical setting as 'phonologico-syntactic'

SLI. Children with SLI were assessed on the same battery of standardized tests of language as the ASD group and they were administrated both experimental tasks (NWR and SR) and RPM, just like the children with ASD.

Eighty-four TD children aged 4 to 12 were recruited from schools in the Tours area. The group was divided in two subgroups of 42 younger children aged 4 to 5 years, TD4-5 (*M* = 5;0, *SD* = 7) and 42 children aged 6 to 12 year (TD6-12), age-matched with the ASD and SLI groups (*M* = 8;0, *SD* = 19). The younger subgroup was assessed via the same protocol as the children with ASD (standardized tests of language abilities, experimental tasks and RPM). Regarding standardized tests, the children in the TD4-5 group were administrated the BILO battery for pre-school children. Six children in this group were assessed only on the experimental tasks and RPM. The older subgroup (TD6-12) was assessed only on SR and on RPM. We excluded any children with learning problem or who were being followed by a SLP or a psychologist. The TD6-12 children were not administered the NWR task because it was assumed that performance on phonology at age 6-12 would have been at ceiling. Further evidence for this hypothesis was the fact that TD children at age 5 typically show ceiling performance on the NWR task.

The data for both the SLI and the TD6-12 groups came from the BiLaD (Bilingual Language Development) ANR/DFG project (2012-2016) aiming to identify SLI in monolingual and bilingual children. In conclusion, the present study focusing on monolingual children includes 147 participants: 37 children with ASD; 26 children with SLI and 84 TD children. The same study on bilingual children will be presented in Part III of the present work.

4.7 Method

4.7.1 Data collection, procedure and coding

The five tasks of standardized language measures, the two experimental tasks of NWR and SR and RPM were part of our research protocol. Each child was tested individually in a quiet room at the hospital that was familiar (when possible) to the child. All sessions were coordinated with the teachers and the teacher’s aids so that the children would not miss classes or therapy. All children were told they were going to play some games. They were told that we could stop whenever they wanted to go back to class or they got tired. Upon completion of each task the child received a sticker as a prize, which he/she stuck on a paper with a grid containing cells for the number of individual tasks in the protocol. The children were assessed in two or more sessions, with a maximum of five (*M* = 2.01; SD = 0.4), whose duration varied from ten to thirty minutes depending on the child's level of attention and participation. Mean session duration was around 20-25 minutes. All sessions were audio-recorded with high quality microphones and then transcribed and coded. The order of the tasks proposed in each session could vary to accommodate the child's fatigue. However we followed some general rules. RPM was always given as the first task in a session. This task took between 4 – 12 minutes, depending on the child attention and motivation. In this way we tried to prevent any effect of fatigue. Moreover, both receptive tasks (vocabulary and morphosyntax) were proposed after expressive tasks in order to prevent word-learning effects or structure-learning effects. All tasks, except the ELO receptive vocabulary task and RPM were administered via computer, and all stimuli were pre-recorded by French native speakers. If a SLP had evaluated a child on one of the five tasks from our protocol within the previous six months, we took the results of this assessment and did not propose the test again. This choice was made in order to prevent possible learning effects.

The NWR task used a PowerPoint presentation with pre-recorded audio stimuli. In order to make the test attractive, an alien appeared on the screen. The child was told that the alien came from another planet and would like to teach him/her its language. The child was then asked to repeat the nonwords exactly as he/heard them. Nonwords pronounced incorrectly were transcribed in a pre-existing coding file. Each transcription was also blind-checked by an expert phonologist. A nonword was considered incorrect when at least one segment was deleted, added, or replaced. The only exception was voicing substitutions (for example [p] repeated as [b]) and substitution of similar vowels (for example [e] repeated as [ɛ]). As the aim of the task was to assess phonology and not articulation, we also did not count as an error any sound that was systematically substituted by another sound (for example, [t] systematically substituted by [k]). Once the transcription and the checking were completed, each error was coded according to its position in the nonword and its type (deletion, substitution, metathesis, etc.). From this coding, we computed a score based on exact repetition: each nonword was coded as 1 when it was correctly repeated and 0 when it was repeated with at least one error. In this study, we limited our analyses of computational complexity to syllable length and errors on consonant clusters, which were identified as two of the main sources of errors in children with SLI for the NWR task. We will come back to the analysis of these complex features in Chapter 7.

For the SR task, the sentences were pre-recorded by a French native speaker and presented, in pseudo-randomized order, through a PowerPoint presentation along with an interactive image which showed the child his/her progression during the task. The sentences were played only once unless external interruptions occurred. The experimenter was instructed to give no response-contingent feedback, but only general encouragement. Each transcription was also blind-checked by an expert syntactician. Repetitions were scored as correct or incorrect according to three measures: if they were verbatim repetitions of the stimulus sentence (Identical Repetition), if they were grammatical (Grammaticality) and if they preserved the structure targeted in the stimulus sentence (Target Structure) even in presence of other errors (substitutions or omissions). For example, the sentences repeated in (5), (6) and (7) were coded as in Table 10.

(5) Input: *Les chats boivent du lait*

The.pl cats drink.3pl some.masc milk

(Structure: SVO present tense 3P)

Child repetition: Les chats boivent du lait

'The cats are driking milk'

(6) Input: *J'ai vu le chat qui a griffé la vache*

I've seen the.masc.sg. cat who has scratched the.fem.sg. cow

(Structure: Subject relatives)

Child repetition: Le chat il a griffé la vache

'The cat he has scratched the cow'

(7) Input: *Les parents ont rangé les jouets*

The.pl parents have.3pl put+away the.pl toys

(Structure: SVO past tense 3P)

Child repetition: Les parents ont joué les jouets

'The parents played the toys'

TABLE 10. IDENTICAL REPETITION, GRAMMATICALITY AND TARGET STRUCTURE

Repeated sentence	Identical repetition	Grammaticality	Target structure
(5)	1	1	1
(6)	0	1	0
(7)	0	0	1

Key: (1 Correct; 0 = Incorrect)

Sentence in (5) was a verbatim repetition of the given stimulus, for this reason we assigned 1 in the Identical repetition score; moreover, this kind of production were considered also grammatical (1 for Grammaticality score) and target (1 for Target structure score). Production like (6), were not verbatim (0 for Identical repetition score), they were not target since the main structure of the sentence was changed (0 for Target structure score), but they were grammatical (1 for Grammaticality score). Finally repetitions as (7) obtained 0 for Identical repetition score, 0 for Grammaticality score and 1 for Target structure score, since the sentence preserved the structure of the sentence. Detailed coding of all errors children made on sentence repetition was carried out. We will come back to the analysis of these errors in Chapter 7.

4.8 Analysis and statistical procedures

All statistical analyses were conducted with the R-studio version 1.1.423 and SPSS version 21. Since each research question required a specific data analysis, we will introduce our statistical procedure in the methodology paragraph of each chapter of the Results section. For general purposes our analyses were conducted with nonparametric tests, ANOVA by ranks (Kruskal–Wallis test) in order to reveal group effects, the Mann-Whitney test for inter-group comparisons and the Wilcoxon test for intra-group comparisons, associated with Spearman's rank correlations, when we were dealing with non-normal distribution of the data (confirmed by the Shapiro-Wilk test) or outliers. Nonparametric tests are distribution free tests, they can handle ordinal data or ranked data, and they may not be seriously affected by outliers. Nonparametric tests were also used when the analyses involved very small groups (Chapter 6 and 7). When the data distribution was normal our analyses were conducted with Student's t-test (Baayen, 2008). For individual analyses we used the Crawford et al., (2010) Singlims program, which includes a t-test for comparison of a single case to a control population. This allowed us to establish, for each individual with ASD, how similar his/her performance was compared to the control groups. Post-hoc t-tests included Bonferroni correction in order to avoid a type 1 error due to multiple comparisons.

Results

Chapter V
Structural language and nonverbal abilities in monolingual children with ASD: some methodological considerations

5.1 Introduction

In this chapter we will address the first research question: **what are the measures most likely to appropriately assess structural language and cognitive abilities in ASD?** In order to answer this question we will divide this chapter into two parts: the first part (section 5.2) will focus on language abilities, and the second one (section 5.3) on cognitive abilities.

Concerning language abilities, we will first report on general results of the four groups (ASD, SLI, TD4-5 and TD6-12) on all tests evaluating them. Then we will concentrate on a detailed analysis of the performance of the children with ASD in order to verify whether our NWR and SR tasks are better suited than standardized tests for an evaluation of structural language abilities in this population. Based on the conclusions of Chapter 1, we hypothesized that repetition tasks that are linguistically based and targeting computational complexity, both for phonology and morphosyntax, such as the LITMUS-NWR and the LITMUS-SR tests, should be the most reliable tools for evaluating structural language abilities in children with ASD. These tasks should play down the possible influence of other factors (pragmatics, lexical knowledge, use of language in a conversational context, etc.), providing information on formal language abilities not provided by currently available standardized tests and making it possible to detect children with structural language impairment. In order to verify our hypotheses we will seek answers to the following questions:

- ***a.*** *Do children with ASD display spared or even enhanced lexical abilities in relation to structural language abilities?*
- ***b.*** *To what extent does the performance of children with ASD on standardized tasks evaluating phonology depend on previous lexical knowledge? To what extent does their performance on standardized tasks evaluating morphosyntax depend on difficulties meeting task demands or on pragmatic difficulties?*
- ***c.*** *Do LITMUS-NWR and LITMUS-SR provide information on formal language abilities not provided by standardized tests?*

When useful to further prove our hypotheses we will compare the performance of the children with ASD with the performance of the children with SLI on tasks evaluating structural language abilities. This comparison will demonstrate that the children with ASD performed differently from the other clinical group and that the use of standardized tasks may partially obscure fundamental similarities in the structural language difficulties with SLI. In conclusion our results will suggest that using currently available standardized tests for evaluating children with autism on structural language abilities should be reconsidered, in favour of LITMUS-SR (for evaluating morphosyntactic abilities) and LITMUS-NWR (for evaluating phonological abilities).

The second part of this chapter will concentrate on results on tasks measuring cognitive abilities. In Chapter 2 we concluded that NV cognitive tasks, RPM and the PRI of WISC-IV, might be the most reliable tools for evaluating cognitive abilities in children with ASD, compared to FSIQ scores and other scores reporting on specific domains of cognition. These tools can be easily used in research and clinical practice; they provide a full picture of autistic intelligence (strengths and weaknesses), and they can be used throughout the entire spectrum. Moreover, and this is important for the purpose of the present study, being NV tasks, they should not (or at least very remotely) be linked to verbal abilities. This should ensure the foundations for a reliable investigation of the relation between linguistic abilities and cognitive abilities. Although RPM have been argued to be the most suitable test for evaluating autistic intelligence in research, and the PRI has been suggested to be the most appropriate index of the WISC-IV for picturing autistic intelligence in the clinical domain, no study has compared these two measures. We hypothesize that since these tasks have been shown to target specific and different aspects of NV capacities, they can both be useful in a complementary evaluation of NV abilities in children with ASD. Moreover to our knowledge no study has directly verified the actual relation between linguistic abilities and performance on both RPM and the subtests of the PRI. Since in the present study we are arguing for the use of tasks that are nonverbal in nature (in that they should involve minimal or no need for task instructions or for speech skills), we will explore the possible relations between RPM and linguistic abilities and between each subtest of the PRI and linguistic abilities in children with ASD. Because they are NV measures of cognition, RPM and the subtests of the PRI, Block Design, Matrix Reasoning and Picture Concepts, should in principle be unrelated to language. We hypothesize, however, in line with previous studies targeting the TD population, that the Picture Concepts subtask, due to its nature, will be strongly related to language abilities.

In order to verify our hypotheses we will seek answers to the following questions:

- ***d.*** *Do children with ASD display better performance on RPM and the PRI than on the FSIQ and all the other indices of the WISC-IV?*
- ***e.*** *Do RPM and the PRI subtests evaluate different NV abilities or are they strongly correlated?*
- ***f.*** *To what extent are RPM and the subtests of the PRI (Block Design, Matrix Reasoning, Picture Concepts) related to language abilities?*

5.2 Language abilities

5.2.1 Methods, participants and data analysis

For the analysis of language abilities we will present the results as follows. The first section will briefly present the results of all four groups (ASD, SLI, TD4-5 and TD6-12) on all measures of language abilities; the second section will concentrate on the comparison between the performance of the ASD group on the standardized language tests and on the experimental tasks (SR and NWR). When useful to further validate our results we will compare the results of the ASD group with the results of the SLI group.

When the ASD group is compared to the children with SLI, and to the children in the TD4-5 or TD6-12 groups, parametric tests will be used, due to the high number of participants in each group. For the analysis of the ASD group, non-parametric tests with will be used, due to the presence of some outliers and a non-normal distribution of the data (confirmed by the Shapiro-Wilk test).

5.2.2 Global results

The children were evaluated on language abilities via a battery of standardized language tasks including vocabulary in reception (VocR) and production (VocP), morphosyntax in reception (MorsynR) and production (MorsynP), phonology (Phono) and the two experimental tasks, LITMUS-NWR, evaluating phonological abilities, and LITMUS-SR, evaluating morphosyntactic abilities.

Table 11 reports group results on all tasks. The cut-off score for language impairment on standardized tests was established at -1.25 SD (see section 4.5.1.1), while on NWR it was 77 % of correct repetition and for SR it was 78% of correct repetition (see section 4.5.1.2). Looking at group performance in Table 11, we can see that while the ASD and the SLI groups displayed similar performance, with scores below cut-offs on tasks of structural language abilities and spared performance on tasks of lexical abilities, the TD groups were more homogeneous, and, as expected, performed within norms on all tasks.

Considering minimum and maximum values, performance both within and below norms was found in all four groups. For the TD4-5 group low scores were due to some children (n = 7) who showed isolated low performance (see Table 11). None of these children performed below norms on more than one task. For the TD6-12 group, the low score on SR was due to one child who performed below threshold on the test. Comparing the ASD group to the two TD groups, the TD4-5 group performed significantly better on almost every measure: VocP ($t(70) = -4.86$, $p <. 001$), Phono ($t(68) = -3.44$ $p = .001$), MorsynP ($t(67) = - 5.60$, $p < .001$), MorsynR ($t(70) = - 5.82$, $p < .001$) and SR ($t(77) = -3.80$, $p < .001$). VocR and NWR represented the only exceptions. Significantly higher performance was found in the TD6-12 group on the only available measure: SR ($t(77) = -6.18$, $p < .001$). Recall that children in the TD6-12 group were not administered either the standardized tests or the NWR task (see section 4.6.).

Focusing on the ASD and the SLI groups, global results show that both groups displayed heterogeneous performance between tasks evaluating lexical knowledge and tasks assessing structural language abilities. On average, performance in each of these groups was in the normal range on the two Voc tasks and was below norms on all the other tasks (with the exception of MorsynR in the SLI group). The ASD and SLI groups did not differ significantly from each other on any task, except for the two tests evaluating phonology, Phono ($t(58) = 2.32$, $p = < .023$) and NWR ($t(61) = 3.40$, $p = .001$), with the SLI group performing much lower than the ASD group.

If we limited our analysis to these group results, we could conclude that the children with ASD, as a group, tended to perform worse than both the younger and age-matched TD children on almost every measure, while they seemed to show a profile similar to the one displayed by the children with SLI. This profile consisted of spared lexical abilities and impaired structural language abilities. The only exception was represented by phonology where impairment in ASD was less severe than in SLI.

However, there are two main problems with these conclusions: (1) if we look at minimum and maximum values for each task, we can see that not every child with ASD performed within the norms on lexical tasks, and not every child with ASD performed below norms on tasks assessing structural language abilities. It is, then, necessary to look at individual performance in order to gain insights into the heterogeneity of the results. (2) Following previous considerations about the possible influence of other factors (difficulties in using language in context, pragmatics, lexical knowledge) on the performance of children with ASD on standardised tasks evaluating structural language abilities (Chapter 1), the question arises as to what extent the results of the ASD group on these tests represent a truthful picture of their real language capacities. In order to answer this question we will further analyse the performance of our group of children with ASD, comparing their results on the standardized tests and on the experimental tasks (SR and NWR).

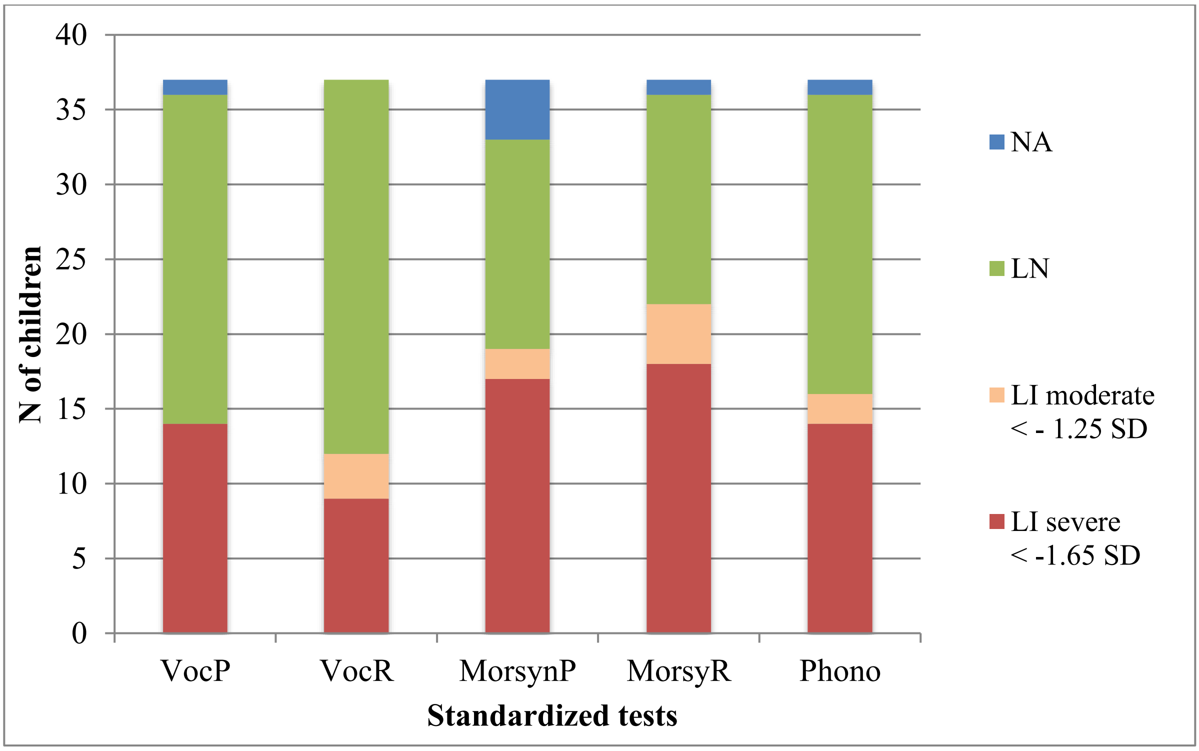
TABLE 11. GROUP PERFORMANCE ON STANDARDIZED TESTS AND ON EXPERIMENTAL TASKS (MEAN, (SD), RANGE)

		Standardized tests					Experimental tasks	
Groups		VocP (z-score)	VocR (z-score)	MorsynP (z-score)	MorsyR (z-score)	Phono (z-score)	NWR (%)	SR (%)
ASD	Number of children that could complete the task /37	36	37	33	36	36	37	37
	Results	-1.01 (1.5)	-0.93 (2.8)	-1.58 (1.8)	-1.79 (2.2)	-2.37 (5)	85 (*15*)	76 (*26*)
		-5.23 - 1.3	-12.07 – 3.09	-5.62 -1.72	-7.01 – 1.64	-25.98 – 3.32	40-89	7-100
SLI	Number of children that could complete the task /26	26	26	26	26	26	26	26
	Results	-1.23 (1.1)	-0.51 (1.5)	-2.04 (1.3)	-0.84 (1.7)	-5 (3)	63 (*22*)	64 (*26*)
		-5.18 – 0.61	-5.74 – 1.43	-4.51 - -0.21	-3.68 – 1.48	-13 - -0.31	3 – 97	0 - 100
TD4-5	Number of children that could complete the task /42	36	36	36	36	36	42	42
	Results	0.4 (0.6)	-0.1 (1.4)	0.4 (0.8)	0.6 (0.6)	0.5 (0.4)	83 (*11*)	85 (*11*)
		-1.2 – 1.5	-6.1– 1.1	-2.1 – 1.9	-0.8– 2	-0.5 – 1.1	44-98	50-100
TD6-12	Number of children that could complete the task /42							42
	Results							96 (*5*)
								77-100

5.2.3 Group performance and individual performance of children with ASD on standardized tests language

As we anticipated in the previous section, when we look at minimum and maximum values on each task, we can see that individual performance in the ASD group was very heterogeneous, including both impaired and spared performance on all five standardized tasks. These considerations raise the question as to how many children with ASD performed below norms on each one of the five tasks and how severe their impairment was. Adopting a cut-off of -1.25 SD for language impairment and separating children between those who showed no deficit (LN) and those who displayed moderate impairment (from < -1.25 to < -1.65 SD) and severe language impairment (< -1.65 SD) we obtain the following picture (Figure 4).

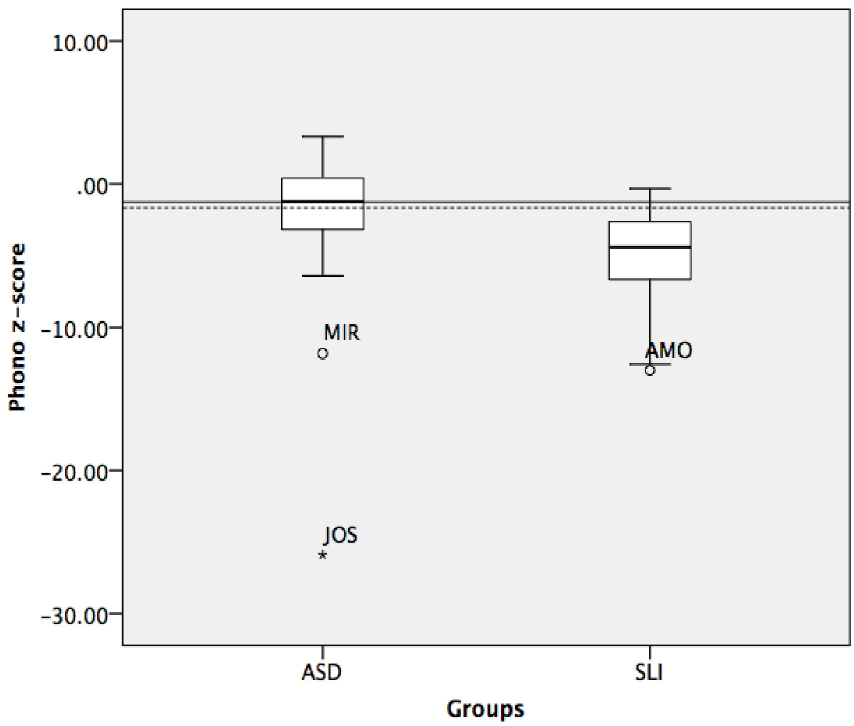
FIGURE 4. STANDARDIZED TESTS: PROPORTION OF CHILDREN WITH ASD WITHIN NORMS AND WITH MODERATE TO SEVERE DIFFICULTIES



Results showed that the total number of children with LI highly varied depending on the linguistic domain being assessed. In general 1/3 of the sample had a LI profile on tasks evaluating vocabulary, while on MorsynR and MorsynP the rate of children with LI was almost twice as big. MorsynR was the task where the largest number of children performed below cut-off for impairment (22/36), followed by MorsynP (19/34), Phono (16/36) and finally by VocP (14/36) and VocR (12/36). Moreover, adopting the cut-off for severe language impairment below -1.65 SD, results indicated that for each task, roughly between 80% to 100% of the children displaying impaired performance showed severe language difficulties.

For Phono, although group results in Table 11 identified this task as the most problematic for children with ASD (the lowest mean rate of performance), the number of children with ASD displaying an LI profile (16/36) was only slightly higher than for vocabulary tasks and it was lower than for both MorsynR and MorsynP. Why, then, was the mean rate of group performance on this task so low? Looking at individual performance we noticed that two children (MIR and JOS) performed so poorly on the Phono task (MIR = -11.83 SD and JOS = -25.98 SD) that their scores lowered the performance of the entire group (Figure 5). If we exclude these *outliers*, the mean rate, although it remains below the norm, goes up to -1.33 SD (compared to -2.37 SD previously).

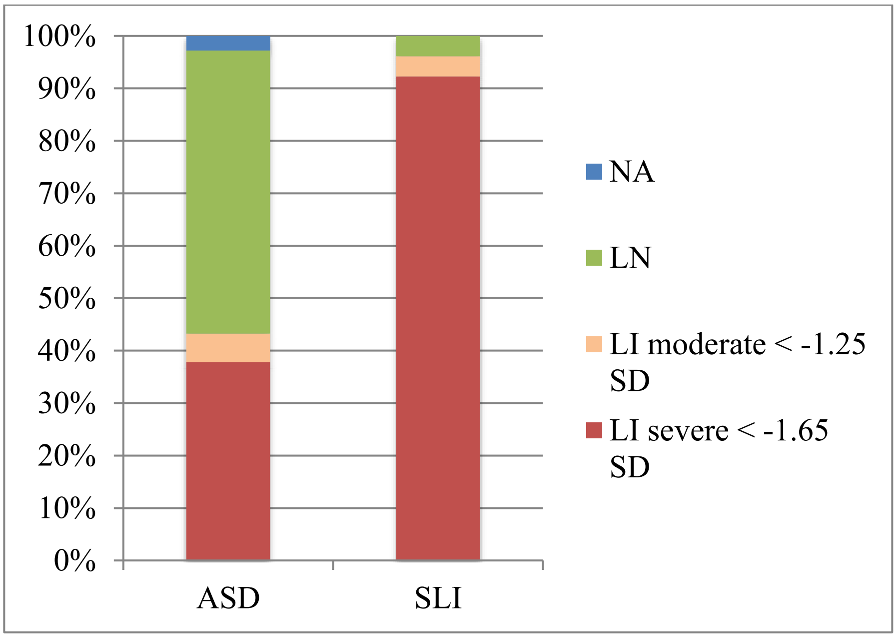
FIGURE 5. ASD AND SLI GROUP PERFORMANCE ON PHONO TEST



Note: The cut-off for moderate impairment was -1.25 SD (indicated by the solid line) and for severe impairment it was -1.65 SD (indicated by the dotted line)

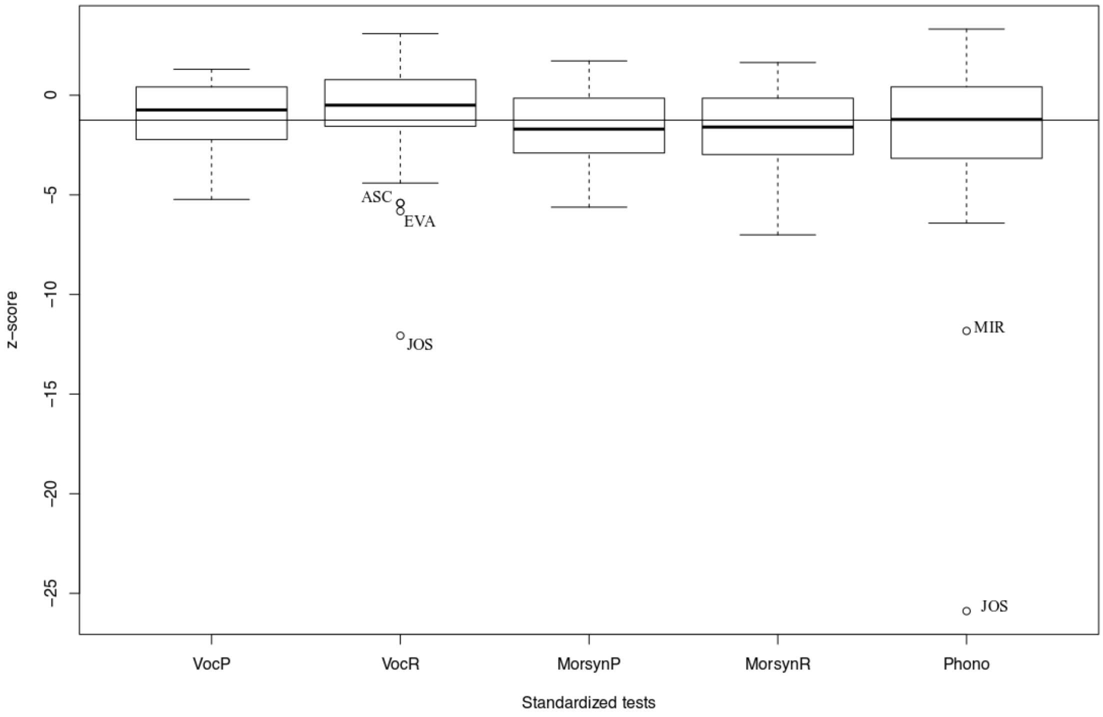
General results also highlighted significant differences between the performance of the children with ASD and the children with SLI on the Phono task, with the SLI group scoring much lower than the ASD group. As can be seen in Figure 6 phonology was impaired in all children with SLI except for one child (25/26), with 24/25 showing severe deficits (< -1.65 SD). In the ASD group, on the other hand, only 16/37 children displayed impairment in Phono, with 14/16 showing severe difficulties.

FIGURE 6. PHONO TASK: PROPORTION OF CHILDREN WITH ASD AND SLI WITHIN NORMS AND WITH MODERATE TO SEVERE DIFFICULTIES



For VocR three children in the ASD group (ASC, EVA and JOS) also performed very poorly. Nonetheless, when excluded, group performance did not change significantly (*M* = -0.63), since roughly 68% of the children in the ASD group still performed within norms (Figure 7).

FIGURE 7. BOXPLOT OF ASD GROUP PERFORMANCE ON STANDARDIZED TESTS



Note: The cut-off for performance below norm was -1.25 SD (indicated by the solid line)

Individual profiles of performance on each one of the five standardized tests showed that the evaluation of linguistic abilities in multiple domains (vocabulary, morphosyntax and phonology) did not provided a clear picture of language capacities in the ASD population. The children displayed heterogeneous performance not only across lexical and structural language abilities, they also showed further discrepancies from one test to another within the same language domain. A detailed overview of individual performance is shown in Table 12.

TABLE 12. INDIVIDUAL PERFORMANCE OF THE CHILDREN WITH ASD ON ALL FIVE STANDARDIZED TESTS INCLUDED IN THE RESEARCH PROTOCOL

Child	Age (months)	VocP	VocR	MorsynP	MorsynR	Phono
SIM	75	-2,15	-0,63	-1,03	-1,87	-2,16
SCO	81	0,41	-5,41	-1,74	-1,74	-1
RUG	82	-2,54	-1,54	-2,15	-1,37	-1,23
ASC	84	-0,58	-5,41	-3	-2,14	0,46
KEV	85	-2,11	-4,41	-1,8	-2,66	-1,7
VOR	87	-0,3	0,47	0,35	0,39	0,42
MTH	87	-1,97	1,42	-2,9	-5,48	-4,32
YLA	88	-2,45	-1,9	-1,9	-3,15	-1,16

LWA	90	-3,11	-1,42	-4,48	-2,67	-6,42
ODI	90	1,16	0,67	-2,9	-3,16	-2,21
EVA	93	-2,6	-5,82	-1,7	-4,78	-6,42
MON	94	-0,07	1,42	0,3	1,505	-0,11
CUT	96	-0,3	1,42	-0,15	-0,84	0,42
CIP	98	-1,13	-0,61	-2,42	-2,57	-1,5
NUG	100	-2,8	-1,23	-1,21	-3,75	-2,33
ROS	100	0,59	-0,61	-2,2	-1,45	-2,33
NAF	100	-2,34	-2,7		-4,8	-3,17
NOS	100	1,2	1,3	-0,64	-0,42	-0,67
ATE	106	-0,28	0,78	-0,25	1,15	-1,42
EDT	107	-2	-0,25	-3	-1,44	1,67
EPI	111		-3,73		-4,78	
EMP	111	-2	-1,09	-2,7	-1,46	-5,31
TUC	112	1,3	1,71	1,5	0,9	3,32
GOT	114	-0,28	1,71	-0,35	1,64	0,8
ELO	115	-0,52	-1,56	-3,7	-2,80	-0,22
AVI	118	0,42	1,51	0,23	0,95	0,89
MUG	118	0,6	3,09	1,1	0,2	0,8
JOS	118	-5,23	-12,07		-7,01	-25,89
FIZ	119	-4	0,4		-2,57	
GHO	120	-1	-0,5	-1,36	-0,67	-4,11
ADO	130	0,5	0,42	-5	-0,34	-3,44
LPG	130	-0,16	-0,25	-0,03	0	-0,31
MIR	135	-0,9	-0,65	-5,62	-1,9	-11,83
LAT	139	0,96	0,74	0,82	-0,3	-1,2
MOI	139	0,59	1,24	-1,5	-0,3	0,8
YAT	141	-2,3	-1,76	-4,47	-7	
LEC	145	-1,13	0,74	1,72	0,23	0,3

Note: Orange boxes indicate moderate impairment at < -1.25 SD, red boxes indicate severe impairment at < -1.65

Table 12 shows that 3 children performed below norms on all 5 tasks, 11 children systematically showed LN performance and 5 children displayed LI performance on all tasks they were able to complete. Among the remaining 18 children, no clear pattern seemed to emerge. Concerning lexical abilities, 8/18 children showed discrepant performance between VocR and VocP. On structural language abilities, only 6/18 children displayed impaired performance on all 3 tasks (MorsynP, MorsynR and Phono), while the rest systematically showed impaired performance on MorsynR and varied on whether MorsynP and Phono were also impaired or not. No child displayed selective impairment in lexical abilites (both VocP and VocR), while one child (ATE) showed selective impairment on

Phono and one child (MOI) displayed selective impairment on MorsynP. Finally, our results also showed that VocR did not predict language status in our group of children with ASD, being rarely in line with the scores of structural language abilities.

In conclusion, heterogeneity seemed to characterize language abilities in children with ASD and these results clearly highlighted the limitation of using group results for describing language performance in this population. Moreover, the fact that the children displayed such different performance across the tests (sometimes in the same linguistic domains) raises the question as to whether the standardized tasks we used for evaluating language abilities in children with ASD may have had an influence on whether a child was identified as being language impaired or not. We will now specifically explore each of the three research questions listed in section 5.1.

5.2.3.1 Lexical abilities and structural language abilities

Question 1a: *Do children with ASD display spared or even enhanced lexical abilities in relation to structural language abilities?*

Previous studies suggested that lexical ability involving knowledge of individual words can be a spared, or even enhanced, domain compared to structural language abilities in children with autism (for an overview see Walenski et al., 2006; 2008) and that some tests, such as the EVIP (the French version of the Peabody, PPVT) may actually overestimate linguistic abilities in children with autism (Mottron, 2004). As we already said in Chapter 1, however, none of these studies compared the performance of children with ASD on both vocabulary tasks and tasks evaluating structural language abilities. In this section we aim to investigate these proposals.

Looking back at Table 11, group results showed that the children with ASD displayed heterogeneous performance across tasks evaluating lexical knowledge and tasks assessing structural language abilities. Intra-group comparisons (adjusted with Bonferroni correction $p < .005$) showed that performance on VocR was significantly higher than on MorsynR ($Z = -2.854$, $p = .004$), while a tendency toward significance was found between VocR and MorsynP ($Z = -2.749$, $p = .006$) and between VocR and Phono ($Z = -2.573$, $p = .010$). These results seem to be in line with what has been suggested by Walenski (2006; 2008). Children with ASD generally displayed, as a group, good performance on VocR and VocP and impaired performance on structural language abilites.

However, looking at individual performance, we saw that some children showed severe impairment on vocabulary tasks as well. Individual results (Table 12) showed that 18/37 children displayed performance on Voc below cut-offs for moderate impairment, < -1.25 SD (*n* = 1), and severe impairment < -1.65 SD (*n* = 17) on one or both tasks. These considerations call into question the statement regarding spared lexical abilities in ASD. Moreover, these findings raise the question of whether children with ASD exhibit enhanced lexical abilities in relation to structural language capacities. We saw in Figure 4 that 1/3 of the children with ASD performed below norms on vocabulary tasks, while the number of children displaying impairment on structural language abilities was almost twice as big, at least on MorsynP and MorsynR. Could vocabulary be affected independently from structural language in ASD? While among the ASD group, all twelve children who performed below norms on VocR systematically showed impaired performance on both MorsynR and MorsynP (when scores were available), for VocP ten children showed impaired performance on morphosyntacictic tasks and two children (SIM and NUG) displayed a selective impairment on MorsynR. Despite these discrepancies, no child obtained an impaired score on one or both vocabulary task(s) and displayed normal performance on both morphosyntactic tasks. For phonology the question was more complex, with 17/34 children (three did not complete the task) showing discrepant performance between VocR and Phono and 10/34 displaying discrepant performance between VocP and Phono, in both directions. In short, in our group of children with ASD, vocabulary did not seem to be affected on its own, but when affected, at least MorsynR was systematically affected as well. This did not held true for MorsynP and phonology, where the children with ASD could display LN performance despite impaired score on vocabulary tasks.

It remains to be seen as a corollary result whether children with ASD really display selective enhanced abilities on the EVIP task, following Mottron's (2004) hypothesis. One way to answer this question is to compare the performance of the children with ASD on the EVIP task with their performance on both another task evaluating VocR and tasks of structural language abilities. Our study allows these analyses. As explained in Chapter 4 (section 4.5.1.1.), half of our group was assessed on VocR through the use of the ELO test and half via the EVIP. After dividing our total sample into two subgroups based on the test they were assessed with, we obtained the following results: the ELO group (*n* = 19) performed in the impaired range (*M*= -1.65, SD = 3.3), while the EVIP group (*n* = 18) performed in the norms (*M* = -0.71, SD = 2.2). A significantly higher performance was found for the EVIP group ($U(36) = 77$, $p = .004$, $r = -.47$). These two groups did not differ on cognitive abilities (measured with both RPM and the PRI) and severity of autism symptoms (measured with ADOS, ECA-R and CARS). Moreover, when compared on performance on standardized tests of language abilities, 3/18 children in the EVIP group showed enhanced score (LN profile) on VocR compared to their scores on structural language tasks (LI profile), while 6/18 children displayed the same results in the ELO group. These results contradict the hypothesis that children with autism exhibit enhanced performance on the EVIP task, in general and in relation to structural language abilities.

To sum up:

Do children with ASD display spared and/or enhanced lexical abilities in relation to structural language abilities?

Findings reported in the literature (Walenski, 2006) predicted that vocabulary is a domain of spared or even enhanced language abilities in children with ASD, compared to structural language. Our results seem to contradict these hypotheses. The fact that over 1/3 of the group had impaired performance on vocabulary tasks and that roughly half of the group (n = 17) displayed severe impairment on one or both vocabulary tasks indicate that lexical abilities were not generally spared in our population sample. Regarding the assumption that lexical abilities are enhanced in relation to structural language abilities, our results only partially confirmed this hypothesis. All children systematically showing impaired performance on vocabulary tasks also displayed scores below cut-offs on MorsynR. However, this did not hold true for MorsynP or Phono, where some children displayed LN performance despite LI performance on lexical tasks. These results are partially in line with findings in Sukenik (2017), who found that children with ASD did not show impaired vocabulary abilities without also having impaired morphosyntactic skills. A corollary result concerned the possibility that performance of children with ASD on VocR could be strongly dependent on the test used for assessment. Our findings did not confirm previous results by Mottron (2004). Although the children evaluated via the EVIP task, as a group, performed better than the children assessed via the ELO task, no significant difference was found for enhanced lexical abilities evaluated via EVIP in relation to structural language abilities.

5.2.3.2 Factors that can intervene in the performance of children with ASD on structural language tasks

Question 1b: *To what extent does the performance of children with ASD on standardized tasks evaluating phonology depend on previous lexical knowledge and on working memory abilities? To what extent does their performance on standardized tasks evaluating morphosyntax depend on difficulties meeting task demands or on pragmatic difficulties?*

Recent studies have suggested that low performance in children with autism on language tests targeting morphosyntax may be the result of a misunderstanding of the pragmatics of the testing situation, while tasks of repetition of real words may engage the use of a pre-existing lexical knowledge, which in relation to word familiarity or frequency effect can lead to biased performance. Moreover, it has been suggested (Kjelgaard & Tager-Flusberg, 2001) that tasks of word-repetition could be dependent on memory abilities, yielding scores related to the capacity of retaining words rather than to phonological complexity itself. This section will investigate these hypotheses. Moreover, we will compare the performance of the children with ASD with that of the children with SLI. These comparisons will demostrate that the children with ASD performed differently from this other clinical group on tasks of morphosynatx and phonology and that these differences might be due to biased performance by the children with ASD on standardized tests.

Starting with phonology we asked whether the performance of our group of children with ASD on the BILO word repetition task was related to lexical knowledge. The BILO task was composed of 40 words of increasing length and complexity. As explained by the authors of the test, one of the main characteristics of the words included in the task was that some of these words were unfamiliar to children, and thus served as quasi-nonwords. We would expect, then, that the presence of words that were generally unknown to children would decrease the possible relation between repetition rates and lexical knowledge. However, as shown in Table 13, when we searched for possible correlations between Phono and lexical abilities both in comprehension (VocR) and in production (VocP), results showed that Phono was strongly correlated with both VocR and VocP in the children with ASD. In the SLI group, we did not find such a correlation.

TABLE 13. PEARSON'S CORRELATION BETWEEN THE TWO VOCABULARY TASKS AND PHONO

	ASD (n = 34)	**SLI (n = 26)**
VocP / Phono	r_s = .510, p = .002	r_s = .214, p = .293
VocR / Phono	r_s = .543, p = .001	r_s = .051, p = .803

In order to verify that enhanced working memory abilities were not responsible for the performance on the repetition tasks, we ran correlation analyses between Working Memory Index (WMI of WISC-IV) scores and scores on the Phono task. Results showed a very strong correlation between these two measures (r_s = .815; p < .001), indicating that performance on word repetition task was related to WM abilities of children with ASD.

These results are in line with previous findings in the literature showing a strong correlation between lexical knowledge and general performance of children with ASD on word repetition tasks (Coady & Evans, 2008). In this vein, our results seem to suggest that children with ASD may rely more heavily on lexical knowledge when they were asked to repeat quasi-nonwords than children with SLI. However, can we really conclude this? as already seen in section 5.2.3, almost half of the children in the ASD group showed discrepant performance between lexical tasks and the Phono task. These results suggest that even if the correlation was strong and showed a general group tendency, the children with ASD could display heterogeneous performance that was obscured by group results. Notably, if lexical abilities play a large role on phonological tasks of word repetition we could suppose that children displaying good lexical skills would perform correctly on word repetition task, which was not systematically the case. For the SLI group we did not find such correlations. These results could be due to the fact that almost the totality of the group performed below norms on phonology, so that no statistical correlation could emerge because of floor effects.

Moving to morphosyntax, we asked whether low performance on standardized tasks displayed by our group of children with ASD depended on difficulties meeting task demands and on pragmatic difficulties. In order to answer these questions, we verified whether possible misunderstandings due to the complexity of the structure of the tasks and to involvement of pragmatic abilities could be spotted. We will focus our attention firstly on MorsynR, and than on MorsynP.

Receptive morphosyntax (MorsynR) was evaluated through a sentence-picture matching task, the BILO "oral comprehension" subtest, which consists of 22 sentences evaluating both specific syntactic structures (n = 14), e.g. clitics, passives, wh-question, etc. as in (8), and inferential statements (n = 8) whose interpretation can only be deduced from pragmatics, world knowledge and a careful observation of the four pictures being proposed, as shown in (9).

Grammatical item

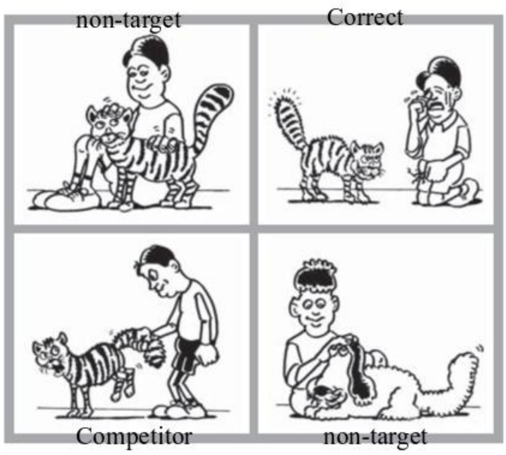
(8) La petite fille lui brosse les cheveux

'*The little girl combs her hair*'

Inferential item

(9) Le chat dont j'ai tiré la queue m'a griffé

'The cat whose tail I pulled scratched me'



As stated in the BILO manual, the choice for the correct answer of item in (8), which is picture number three, depends on the processing of the dative clitic pronoun. If children failed on these types of items, we should consider their answers as *incorrect* since they implied a problem with linguistic processing of the sentence. These conclusions were drawn from the test's standardisation by a group of 782 TD children, aged 7- to 14, who did not show any preference for the incorrect pictures, independently from their age. On the other hand, the processing of items such the one in (9) was possible if the participant understood all the events of the statement in their chronological order and inferred the correct interpretation from the value of the present perfect tense of the verbs. The authors argued that while the choice of pictures 1 and 4 should be considered as *non-target*, because they are not related in any way to the stimulus, picture 3 could be considered as a *valid competitor* of the correct answer. This picture focalises on the first event of the sentence (*I pulled the cat's tail*), even though it does not integrate the second event of the sentence (*the cat scratched me after I pulled his tail*). Crucially these labels were assigned due to the fact that while no TD child chose the non-target pictures (1 and 4), the choice of the competitor picture was very frequent and decreased with participant age, moving from 42% at age 7 to 16% at age 14.

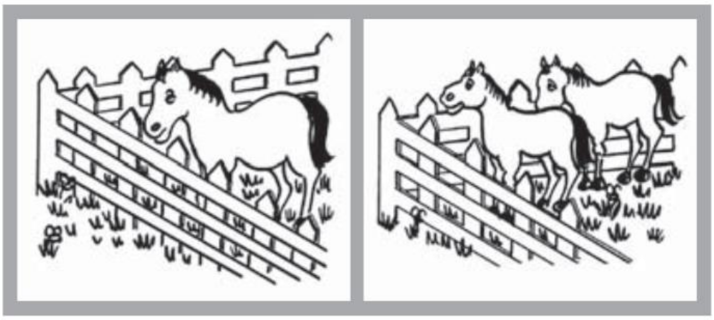
The question was how the children with ASD behaved with respect to grammatical items and inferential items. Results showed that as a group the children with ASD showed higher rate of errors on inferential statements (64% errors on the total number of inferential items) than on grammatical statements (53% errors). Since inferential language is part of pragmatic knowledge (see section 1.2.1) and is known to be generally problematic for children with ASD (Norbury & Bishop, 2002), the high error rate on these items was not surprising. Nonetheless, what was really interesting was the fact that among the total number of errors on inferential items, almost half of them (83/189) were due to the choice of non-target pictures, a result that was not found in TD children. Moreover the behaviour of some children on these types of non-target answers was unusual for the TD population. Pictures in this particular task were very hard to interpret and sometimes they were very distracting for many children with ASD, due also to the level of detail included in the pictures (e.g. picture in (9)). Half of our population sample employed either repetitive strategies, choosing pictures on the basis of their position on the screen (e.g. one child only chose pictures that were on the bottom-left corner of the laptop) or following recurring sequences (they completed the task choosing pictures in the same order 1, 2, 3, 4). Three children clearly stated that they picked the pictures on the basis of their visual preferences (e.g. for the picture in (9), one child stated that he chose the fourth picture because he liked dogs better than cats).

In addition, the children with ASD displayed a high rate of errors on grammatical items. Some of the erroneous answers given by the children with ASD (56/274) were due to the choice of incorrect pictures. These types of errors were not found in TD children (BILO manual) indicating that the children with ASD showed very peculiar behaviours on this task. In contrast, although the children with SLI generally displayed high error rate for both grammatical (32% of errors on total grammatical items) and inferential items (20% of errors on total inferential items), they behaved "typically" in that they did not choose the non-target pictures in the case of inferential statements

For all these reasons, MorsynR does not appear as a reliable tool for evaluating morphosyntactic abilities of children with ASD. The risk of biased performance is too high.

Expressive morphosyntax was evaluated with a sentence completion test, again taken from BILO, consisting of 25 items, in which children had to complete a sentence on the basis of a picture and a verbal stimulus, as shown in (10).

(10)



As the first picture appeared on the screen, a pre-recorded voice said:

Stimulus 1: Ici, il y a un seul cheval

'*here there is only one horse*',

At this point the second picture, with two horses, appeared on the screen and the pre-recorded voice started a new sentence, which the child had to complete as follows:

Stimulus 2: et là, il y a...

'and here there are ...

Expected answer: …deux chevaux.

'two horses'

In general the test targeted gender agreement (both on adjectives and past participles), number agreement (both on nouns and verbs), verb tense, clitics and passives.

Three types of errors were found:

<u>Type 1</u>: Grammatical errors, as in (11):

(11) Target: Ici la fille a un chapeau, là les garçons *ont des chapeaux*

'Here the-sing girl has a-sing hat, there the-plu boys have some-plu hats'

(e.g.) MTH: *… là les garçons… a des chapeaux ‘there the boys… has some hats’

Type 2: Answer related to the picture but not corresponding to the target answer, as in (12)

(12) Target: Ici le chien boit son lait, là les chiens *boivent leur lait*

‘Here the-sing dog drinks his-sing milk, there the-plu dogs drink-plu their milk’

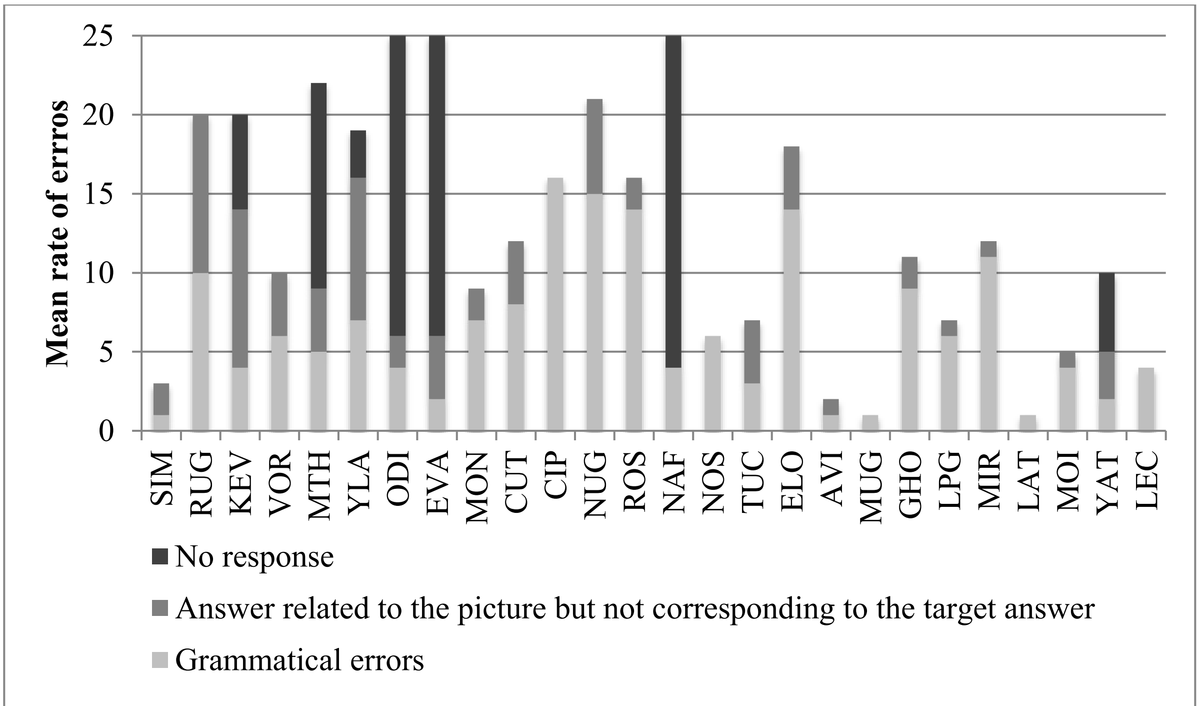
(e.g.) YLA: … là les chiens… aussi

‘there the dogs… also’

Type 3: No response

Figure 8 shows the total number of errors per subject for each of these answer types. In this graph, children are presented by age (from the youngest to the oldest). The graph presents the results of 26/37 children (4 children could not perform the MorsynP task, and for 7 we did not have detailed reports of their performance since the test was administered in a previous session by a speech-language pathologist (see section 4.5.1.1)).

FIGURE 8. TOTAL NUMBER OF ERRORS PER CHILD ON THE MORSYNP TASK



The children with ASD made a high number of errors, which amounted to half of the total number of answers (345/650). Half (*n* = 183) of the total number of errors were morphosyntactic in nature. The remaining errors were of type 2, answer related to the picture but not corresponding to the target answer (*n* = 76), and type 3, no response errors (*n* = 86).

While the total number of errors made by each subject was independent from their chronological age (r_s = .146, p = .417), the error types seemed to be related to their age. Notably, type 2 and type 3 errors were more frequent in younger children than in older ones. No response errors (type 3) seemed to be due to difficulties meeting task demands. The co-occurrence of multiple stimuli, audio and visual, and the need to integrate all information, which consisted in memorizing the first sentence, integrating the new information given by the second picture and finally finding the right answer to complete the second sentence, represented a real challenge for many children with ASD, and younger children seemed to be more sensitive to these types of demands. Despite the high rate, the no response strategy was adopted by a restricted number of young children (7/26). It might be that the co-occurrence of multiple stimuli (visual and verbal) was too distracting for some children with ASD so they ended up not answering the stimuli. Moreover 4/7 children (MTH, ODI, EVA, NAF) were more affected than the others, with roughly 60-80% of their answers being no responses *vs* 20-30% in the other three children.

In contrast, answers like (12) were found almost all over the sample (20/26 children). This type of error was, again, more frequent in younger children than in older ones. Moreover, when we checked for error typology, we noticed that the few errors displayed by older children with ASD were always made on the same stimuli. At some point in the task, the children were asked to switch from 3rd person subjects to 1st person subjects. Two sentences involved such a change of perspective which required the child to identify with one of the characters presented in the picture. Changing perspective in this context is evocative of some characteristics of pretend play, which has been reported to be particularly difficult for children with ASD (Hess, 2006; Tuller et al., 2017; see also section 1.2.1). These stimuli were identified as particularly challenging even for TD children by the authors of the test, who suggested reconsidering these errors when calculating the final score (Khomsi et al., 2007). When these items were excluded, the pattern of results changed significantly for older children who displayed only grammatical errors, but not for younger children, who still showed high rates of type 2 errors. These kinds of answers where produced by each one of the young children on different stimuli. This heterogeneity of answers prevented the computation of a separate grammatical score based only on morphosyntactic errors.

To sum up:

To what extent does the performance of children with ASD on standardized tasks evaluating phonology depend on previous lexical knowledge? To what extent does their performance on standardized tasks evaluating morphosyntax depend on difficulties meeting task demands or on pragmatic difficulties?

Concerning phonology, previous studies evoked the possibility that performance on tasks evaluating phonological abilities via word repetition may depend on previous lexical knowledge (Coady & Evans, 2008) and on WM abilities (Kjelgaard & Tager-Flusberg, 2001). Our results showed that performance on the BILO word-repetition task, although generally correlated with lexical knowledge, was not systematically in line with lexical abilities of children with ASD. While there was a strong group correlation, half of the ASD group showed discrepant performance between lexical and phonological abilities. Moreover, the Phono task was very strongly correlated with WM abilities, suggesting that the length of the stimuli may have played a role in the performance of the children with ASD. We cannot strongly conclude, then, that the word-repetition task of the BILO battery can be either a good or a bad predictor of phonological abilities in children with ASD. We will compare this test with the LITMUS-NWR task in the next section, asking whether NWR is a more reliable measure of phonological abilities in autism.

For morphosyntax, although MorsynR and MorsynP were supposed to concentrate on the evaluation of morphosyntactic constructions, the number of other abilities required to perform both tasks, especially pragmatics (items involving inferential interpretation and knowledge of the world), made it very difficult to obtain a pure evaluation of morphosyntactic abilities. Moreover, the children with ASD showed great difficulties in integrating multiple sources/kinds of information, especially those items involving inferential statements (for MorsynR) and all items involving interpretation of pictures (for both Morsyn R and MorsynP). In the next section we will ask whether SR is a more reliable measure of morphosyntactic abilities in autism.

5.2.3.3 LITMUS-NWR and LITMUS-SR as reliable tools for the evaluation of structural language abilities in children with ASD

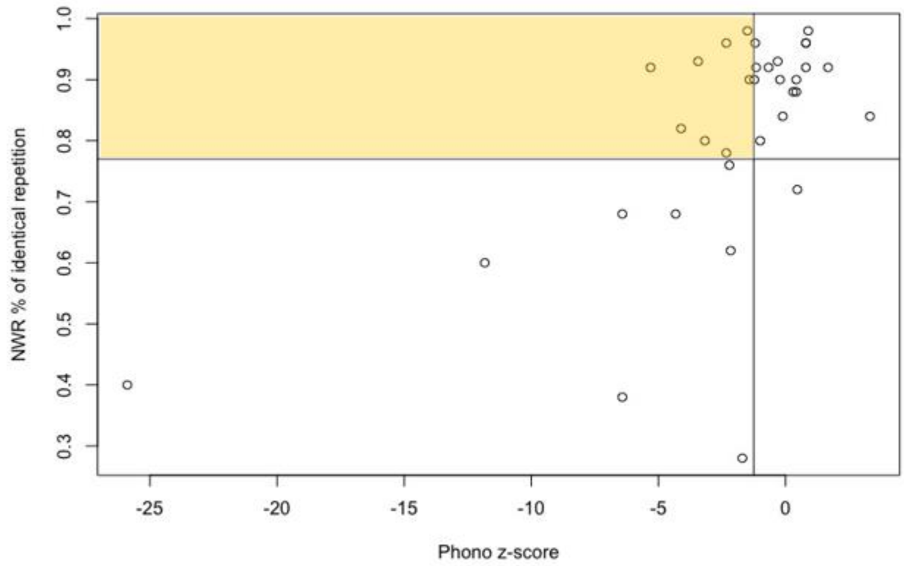
***Question 1c**: Do LITMUS-NWR and LITMUS-SR provide information on formal language abilities not provided by standardized tests?*

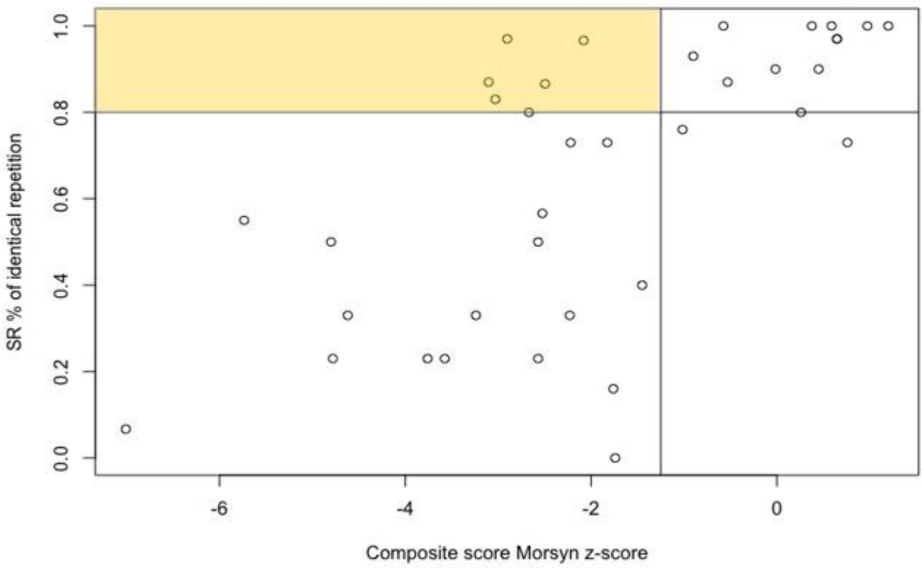
One of the main hypotheses of our study is that LITMUS-NWR and LITMUS-SR, being repetition tasks specifically developed to evaluate computational complexity in phonology and morphosyntax, are the best-suited tools for evaluating structural language abilities in children with ASD. The design of a LITMUS-NWR task should ensure a more controlled evaluation of phonology, since individuals performance cannot depend on possible effects of lexical knowledge (nonwords were conceived to be unrelated to existing words in the child's language) or on WM abilities (nonword were controlled for syllable length) (Kjelgaard et Tager-Flusberg, 2001; dos Santos & Ferré, 2017). Analogously, pragmatic impairment or difficulties in appropriate use of language in context typical of children with ASD should have limited impact on repetition accuracy in SR tasks (Polišenská et al. 2015; Silleresi et al. in press). The design of LITMUS-SR incorporates formal aspects of morphosyntactic processing but it is little constrained by pragmatic features. In order to verify whether both LITMUS-NWR and LITMUS-SR (henceforth NWR and SR) are better suited tools than commonly used standardized tasks for evaluating phonological and morphosyntactic abilities in children with ASD, we will compare performance of our ASD group on these tasks.

Pearson correlation tests were run between scores drawn from the standardized tasks of structural language abilities (morphosyntax and phonology) and repetition tasks (SR and NWR). For phonology, the score of the phonological word repetition task was used, while for NWR we used the total score of identical repetition. For morphosyntactic abilities, we calculated a composite language score combining MorsynP and MorsynR, based on the following rationale. To perform the SR task, participants have to comprehend the sentences in terms of abstract grammatical system representation and then reprocess the linguistic information using their own grammatical and memory systems (Baddeley, 2000). A composite score of MorsyR and MorsynP should represent both of the abilities that are necessary to process and repeat sentences in an SR task. For children that were unable to complete one of these tasks, we used the score that was available. To be sure that no differences would occur with scores of MorsynR and MorsynP when taken separately, we ran following analyses also on the two morphosyntactic scores individually. No difference was found. For SR we used the score of identical repetition.

Strong correlations were found both between the phonology score and NWR (r_s = .490, p = .003) and between the composite morphosyntax score and SR (r_s = .656, p < .001). However, examination of individual performance revealed some discrepancies, both for phonology and for morphosyntax (Figure 9). Although the performance of children on the standardized language tests and the repetition tasks showed similar tendencies, for a high number of children (n =10) with ASD (in the two yellow quadrants of Figure 9) normal structural language abilities were brought to light only by the two repetition tasks (NWR and SR). Eight children (out of 37) performed in the normal range on NWR but not on the Phono task, and six children performed in the normal range on SR but not on the composite score for morphosyntax. Strikingly, the opposite was rarely found. Moreover, four children performed above the threshold on both NWR and SR, while they performed below cut-off on all standardized tasks evaluating structural language abilities.

FIGURE 9. COMPARISONS BETWEEN STANDARDIZED LANGUAGE MEASURES NWR AND SR ON THE ASD GROUP





Note: The vertical lines correspond to the –1.25 SD cut-off for impairment in standardized measures; the horizontal lines correspond to the 77% cut-off for impairment in NWR and the 78% cut-off for impairment in SR (Tuller et al. 2018)

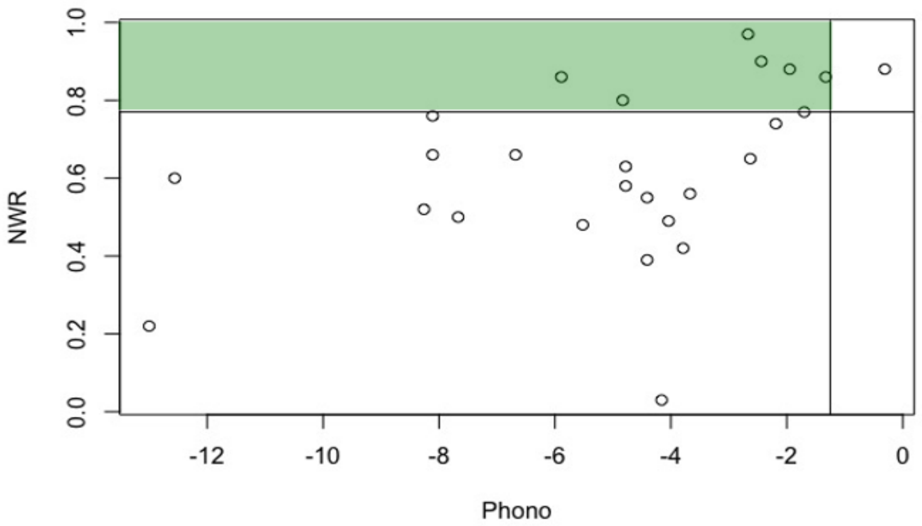
To verify that performance on NWR did not depend on previous lexical knowledge, as hypothesized by the authors of the task (see section 4.5.1.2), we searched for correlations between the two scores of VocP and VocR, and NWR percentages for correct repetition. Strong correlations were found both between VocR and NWR (r_s = .495; p = .002) and between VocP and NWR (r_s = .535; p = .001).

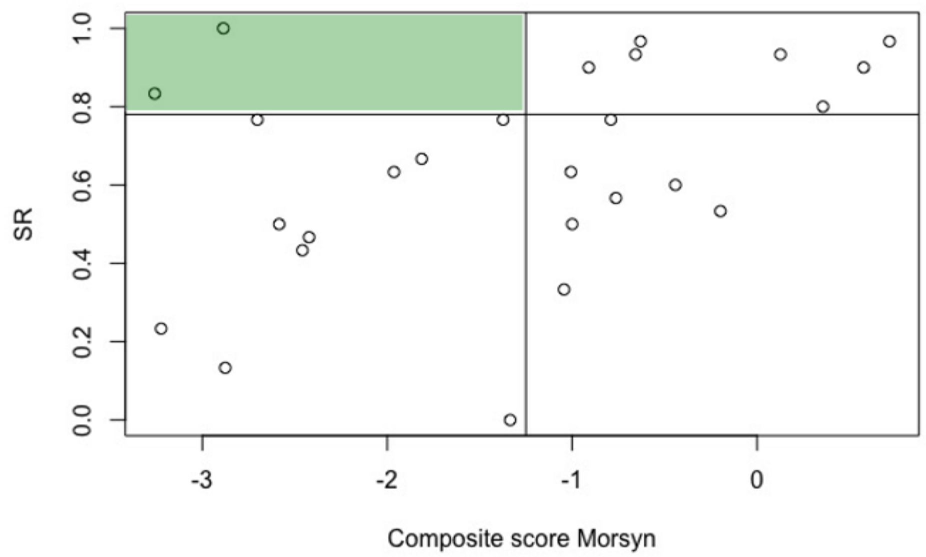
In order to verify that enhanced working memory abilities were not responsible for the performance on the repetition tasks, we ran correlation analyses between the Working Memory Index (WMI WISC-IV) scores and the percentage of identical repetition on NWR and SR. WMI scores were available for only 21/37 children. Results showed a tendency toward a significant correlation between NWR and WMI (r_s = .435; p = .055) and a strong correlation between SR and WMI (r_s = .778; $p < .001$). Nonetheless, when we checked for WM scores of individuals with performance in the norms on SR and on NWR and impaired performance on standardized tasks (the eight children in the yellow quadrants of the first graph of Figure 9 and the six children in the yellow quadrants of the second graph of Figure 9), results showed that five of these children displayed low or very low scores on the WMI of the WISC-IV, indicating that enhanced working memory abilities were not responsible for their high performance on NWR and SR.

We also checked for the tendency of these children to produce echolalic utterances, thus investigating the possible influence of echolalic speech on correct repetition. Taking into account the relevant subscore of the ECA-R (this score was available for all 37 children), six of the children in the yellow quadrants (Figure 9) showed occasional use of echolalic speech, while the other two did not show any tendency to produce echolalic utterances. Moreover, results on the entire group pointed toward a lack of correlation between echolalic speech and NWR (r_s = - .082; p = .648) and a significant negative correlation between echolalic speech and SR (r_s = -.450; p = .008).

Finally, we checked for individual performance on the standardized tests and on NWR / SR in the children with SLI. A strong correlation emerged between the performance on Phono and NWR (r_s = .608; p = .001), while a weak correlation was found between the composite score of Morsyn and SR (r_s = .394; p = .047). Plotting the individual scores of the children in the SLI group (Figure 10) results showed that these children did not display the same pattern of performance as what was found in the ASD group (recall Figure 9), especially concerning SR task. The presence of children performing above threshold in the SR task and below the cut-offs for language impairment on standardized tests was more frequent in the ASD group than in the SLI group.

FIGURE 10. COMPARISONS BETWEEN STANDARDIZED LANGUAGE MEASURES NWR AND SR ON THE SLI GROUP





Note: The vertical lines correspond to the –1.25 SD cut-off for impairment in standardized measures; the horizontal lines correspond to the 77% cut-off for impairment in NWR and the 78% cut-off for impairment in SR (Tuller et al. 2018)

To sum up:

Do LITMUS-NWR and LITMUS-SR provide information on formal language abilities not provided by standardized tests?

Regarding structural language abilities, the children with ASD generally displayed very similar performance on the standardized tests and the two repetition tasks. However, for some of the participants (16% of the group for SR and 22% for NWR), normal language capacities seemed to be revealed only by their performance on NWR and SR, while standardized measures did not provide any additional information on the children's structural language abilities.

We looked at the possible influence of lexical knowledge and working memory abilities on performance in NWR. While both performance on VocR and VocP was strongly correlated with NWR scores, WM abilities emerged as being involved in performing in the BILO word repetition task but not in NWR. These results are in line with the idea that syllable length plays a large role in the BILO word repetition task, but not in the NWR test, for which items were created explicitly with the objective of stemming working memory influence. On the other hand, the results contradict the idea that NWR is not strongly correlated with lexical knowledge. Indeed, as Chiat (2015) hypothesized, it may not be possible to completely isolate nonwords from language knowledge.

Concerning SR we looked for possible influence of WM and echolalia. Echolalia could not be pinpointed as a factor directly responsible for high repetition rates in the children with ASD (in line with Brynskov & Eigsti, 2016), since the strong negative correlation found in our group of children indicated that the children who repeated worse were the ones with strong echolalic use of language. Similarly, WM was not responsible for the results of the children in the yellow quadrants of Figure 9, who showed an enhanced performance on SR in comparison to their scores on morphosyntactic tasks.

5.2.4 Conclusions on language abilities

The purpose of the present section was to verify the assumption that NWR and SR can be the most reliable tools for evaluating structural language abilities in children with ASD, in comparison to standardized tasks evaluating both lexical knowledge and structural language abilities.

Our results confirmed this hypothesis. The group of 37 children with ASD showed better performance on NWR and SR measures than on the other measures of language abilities. In Chapter 1 we predicted that lexical abilities would be excluded from further analyses because they would overestimate the linguistic abilities in children with ASD. Our results showed that it was not the case. Lexical abilities turned out to be severely impaired (< -1.65 SD) in 17 children with ASD and this impairment was sometimes independent from structural language abilities (MorsynP and phonology). However, the heterogeneity of profiles, the fact that children displayed sometimes different results between VocR and VocP and the fact that no clear conclusions could be drawn from lexical abilities in relation to structural language abilities, yielded the exclusion of this task from further analyses.

Concerning phonological abilities, children with ASD exhibited similar performance on both the BILO word repetition task and on NWR. However, since retaining the two measures of phonology could generate redundant effects, only one measure was chosen. As, for some of the participants, normal language capacities seemed to be revealed only by their performance on NWR and given that this task was not related to WM abilities (in contrast to the BILO task), we chose to keep NWR.

For morphosyntactic abilities, we demonstrated that both MorsynR and MorsynP of the BILO were not adapted for returning a faithful picture of morphosyntactic abilities in children with ASD, due to the co-occurrence of multiple factors that intervened in characterizing the performance of these children. Our results confirmed the predictions of Prévost et al. (2017; 2018) and Wittke et al. (2017) who suggested that performance of children with autism on language tests targeting morphosyntax could be the result of a misunderstanding of the pragmatics of the testing situation or of difficulties with appropriate use of language in context. Similarly to what was found for NWR, for some of the participants normal language capacities seemed to be revealed only by their performance on SR. Therefore, results suggested that SR, being controlled for computational complexity and length, and controlled for the possible effects of other abilities (WM, pragmatics, use of language in contexts, inferential statements, etc.) differently from standardized tests, should be the best solution for evaluating morphosyntactic abilities in children with ASD.

5.3 Cognitive abilities

5.3.1 Methods, participants and data analysis

Children's cognitive level was assessed via the RPM task and at least one battery of psychometric evaluation of cognitive abilities. When possible this battery was the WISC-IV. We could gather data on RPM for all 37 children, while data for the WISC-IV were available for 33/37 children, and four children were evaluated via the EDEI-R. When possible for children evaluated via the WISC-IV, we collected scores for FSIQ, the four indices of the battery (PRI, VCI, WMI, PSI) and the three subtests of the PRI (Block

Design, Matrix Reasoning and Picture Concepts). For the four children evaluated via the EDEI-R we collected data for FSIQ, and for the verbal and nonverbal indices.

The first part of the next section will concentrate exclusively on the 33 children assessed via WISC-IV. This choice was made in order to better compare children's performance on the same tasks. The performance of the four children assessed via the EDEI-R task will be briefly discussed in section 5.3.4.

Non-parametric tests with ANOVA by ranks (Kruskal–Wallis test for group effects, the Mann-Whitney test for inter-group comparisons, and the Pearson correlation test) will be used for statistical analysis on the ASD group, due to the small number of participants, the presence of some outliers and a non-normal distribution of the data (confirmed by the Shapiro-Wilk test).

5.3.2 Results

5.3.2.1 Discrepancy between NV cognitive measures and WISC-IV in Autism

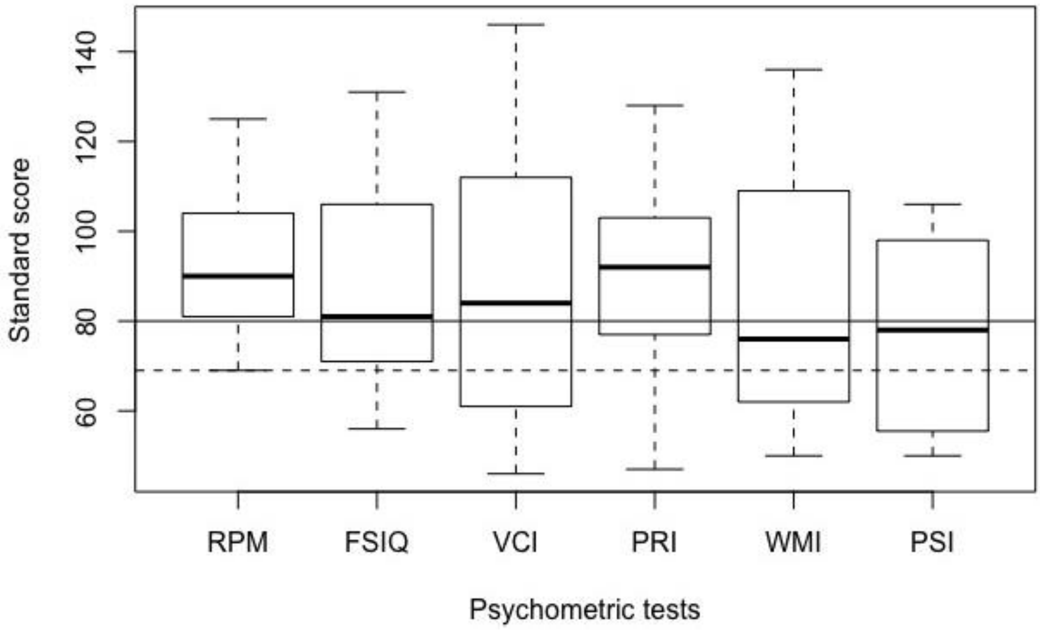
Question 1d: *Do children with ASD display better performance on RPM and the PRI than on the FSIQ and all the other indices of the WISC-IV?*

General results of the ASD group (n = 33) on both RPM and the different indexes of the WISC-IV, as well as the FSIQ score, are reported in Table 14 and Figure 11. The cut-off for impaired abilities was established at < 80 standard scores for every measure. Following the descriptive classification of IQ reported in the WISC-IV manual, the children in our ASD group ranged from moderate intellectual disability to very superior intellectual abilities (47-146 standard score). As can be seen in Table 14, the children displayed performance within the norms for RPM and every index of the WISC except for the PSI. However, significant differences (adjusted with Bonferroni correction $p < .003$) were found between RPM and FSIQ scores ($Z(33) = -5.012$, $p = < .001$), between RPM and the PSI scores ($Z(25) = -4.459$, $p = .001$) and between the PRI and PSI scores ($Z(25) = -3.362$, $p = .001$), always in favour of NV scores. No other significant difference was found.

TABLE 14. ASD GROUP PERFORMANCE ON PSYCHOMETRIC TESTS AND ON RPM, MEAN (SD) AND MINIMUM-MAXIMUM VALUES

Psychometric tests	RPM	WISC IV				
		FSIQ	PRI	VCI	WMI	PSI
Number of children that could complete the task /33	33	33	33	33	21	25
Results (standard score)	93 (14)	86 (21)	91(18)	86 (26)	84 (28)	76 (20)
	69 - 125	56 - 131	47 – 128	46-146	50 – 136	50 – 106

FIGURE 11. ASD GROUP PERFORMANCE ON PSYCHOMETRIC MEASURES OF COGNITIVE ABILITIES

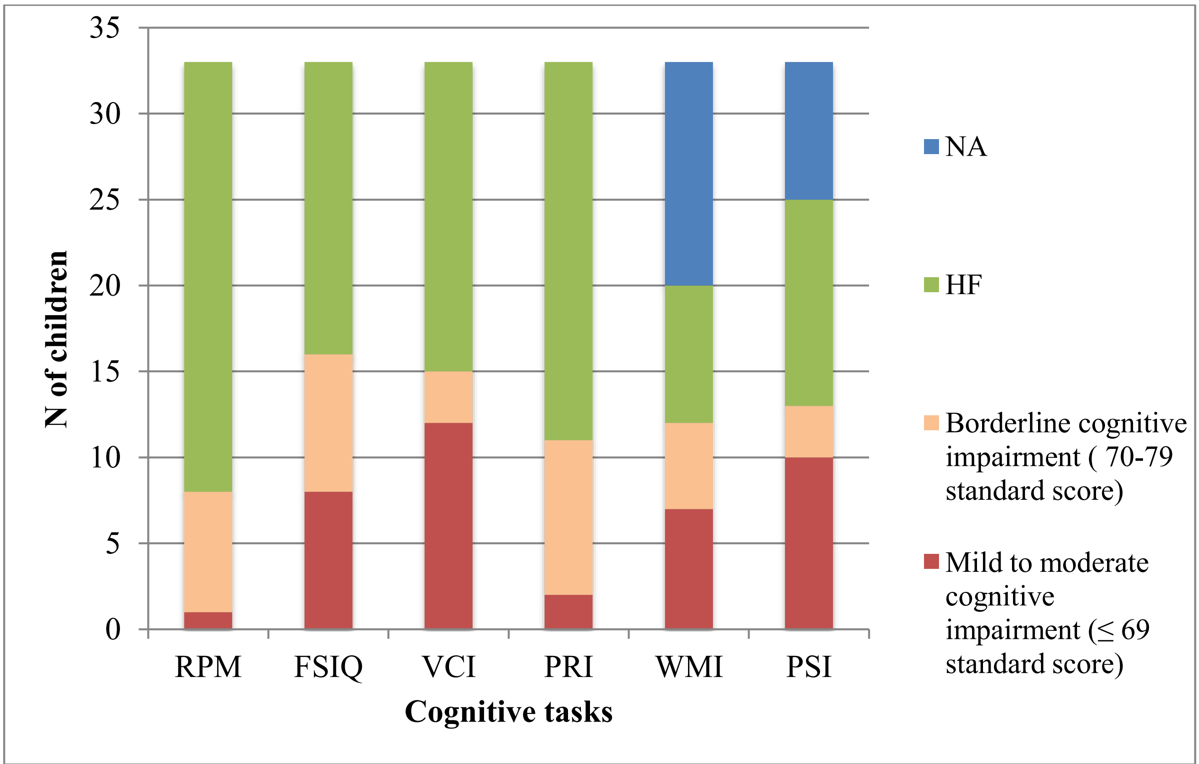


Note: The cut-off for performance below the norm was < 80 standard score (indicated by the solid line); for mild/moderate intellectual impairment it was < 70 standard score (indicated by the dotted line)

The absence of significant differences was due to the great heterogeneity of the results displayed by the children with ASD on every measure, as shown in Figure 11. The children tended to perform better on tasks evaluating nonverbal abilities, RPM and the PRI of the WISC-IV, than on all the other tasks. If we consider minimum and maximum values, performance both within and below norms were found for every index of the WISC-IV and for RPM. The question that arises is how many children displayed impaired performance on all these scores and how severe was their impairment?

Adopting a cut-off of < 80 standard score for cognitive impairment and separating children between those who showed no deficit (≥ 80 standard score, HF), those who displayed a borderline performance (from 70 to 79 standard score) and those displayed mild / moderate cognitive impairment (≤ 69 standard score) we can see that the number of children with cognitive impairment varied depending on the cognitive measures taken into account (Figure 12).

FIGURE 12. PSYCHOMETRIC TESTS: NUMBER OF CHILDREN WITH ASD WITHIN NORMS AND WITH BORDERLINE AND MILD/MODERATE COGNITIVE IMPAIRMENT



In general 25% of the children showed impaired profiles on tasks evaluating NV abilities (8/33 on RPM and 11/33 on the PRI), while on all the other measures the mean rate of children with impaired cognitive performance was close to 50%. Children performed in the impaired range more frequently on the FSIQ (16/33) and the VCI (15/33). For the PSI and the WMI no data could be retrieved for 13 and 8 children respectively. Concerning available data, 12/20 children showed impaired performance on the PSI and 13/25 children performed in the impaired range for the WMI. Moreover, the results indicated that while for the FSIQ, VCI, PSI and WMI the majority of children with ASD showing a deficit in intellectual abilities would be considered mildly to moderately impaired, for RPM and the PRI only three children would end up being categorized this way.

To sum up:

Do children with ASD display better performance on RPM and the PRI than on the FSIQ and all the other indices of the WISC-IV?

Our results are in line with previous findings in the literature that demonstrated that children with ASD tend to perform better on RPM than on the FSIQ and on all indices of WISC-IV (Barbeau et al. 2013; Courchesne et al. 2015; Mayes and Calhoun 2008; Nader et al. 2016). Moreover, our results confirmed previous studies claiming that the PRI of WISC-IV highlights the enhanced cognitive abilities of individuals with ASD in comparison to both the FSIQ and the other indices of the Wechsler scale (Mayes and Calhoun, 2008; Nader et al., 2016).

However, no study has compared these two measures to see if any difference would emerge in the performance of children with ASD. In the next section we will compare all three subtests of the PRI with RPM.

5.3.2.2. RPM and subtests of the PRI (WISC-IV)

Question 1e: *Do RPM and the PRI subtests evaluate different NV abilities or are they strongly correlated?*

No study has directly compared RPM with the PRI of the WISC-IV, although Nader et al. (2016) hypothesized that no difference should be spotted. Yet, even within the PRI, a few studies have reported that children with ASD tend to display a further peak of abilities on the Matrix Reasoning and the Block Design subtests in comparison with the Picture Concepts subtest (Nader et al., 2015; Oliveras-Rentas et al., 2012). As suggested by Houskeeper (2011), these results may be due to the fact that children with ASD show their strengths more consistently on visuospatial tasks (Block Design) and on fluid reasoning tests (Matrix Reasoning), while they should display more difficulties when the task relies on language abilities (Picture Concepts). This section will compare these three measures to RPM in order to see if children with ASD perform similarly on all four tasks.

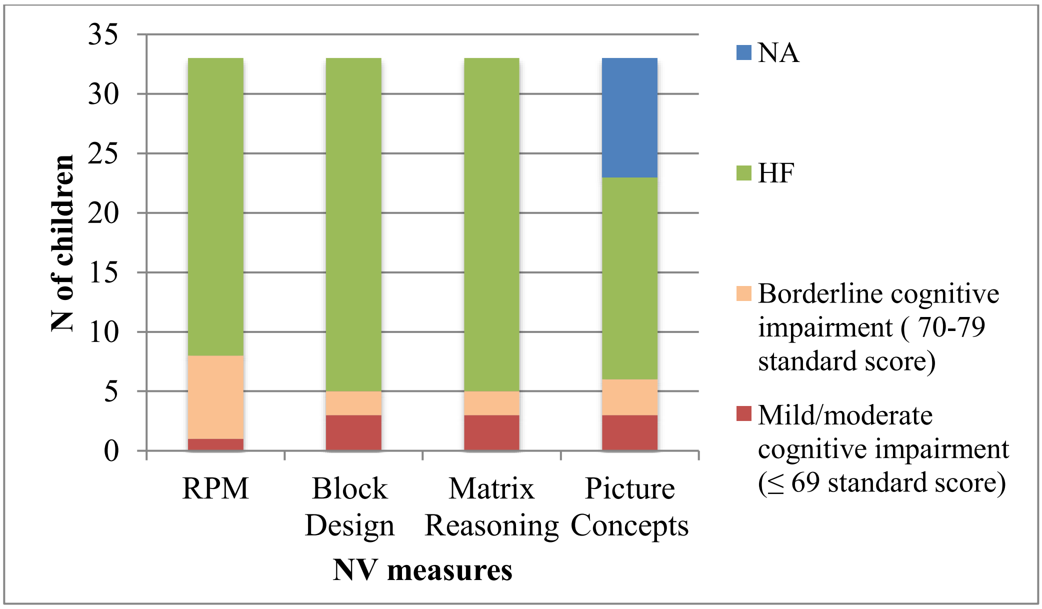
Table 15 reports group performance on the four NV measures. No significant differences between them were found, except for a tendency toward significantly better performance on Block Design compared to Picture Concepts ($Z(33) = -2.332$, $p = .020$).

TABLE 15. ASD GROUP PERFORMANCE ON NV TASKS, MEAN (SD) AND MINIMUM-MAXIMUM VALUES

NV measure	RPM	Block Design	Matrix Reasoning	Picture Concepts
Number of children that could complete the task /33	33	33	33	23
Results (standard score)	93 (14)	96 (18)	94 (16)	87 (17)
	69 -125	55 - 125	60 - 120	55 -115

Figure 13 displays the number of children who performed below the norms on the four NV measures (note that for Picture Concepts data were not available for 10 children due to their inability to perform the task). The cut-off for impaired abilities was established at < 80 standard score; for mild/moderate intellectual deficit it was set at ≤ 69 standard score.

FIGURE 13. NV MEASURES: PROPORTION OF CHILDREN WITH ASD WITHIN NORMS AND WITH BORDERLINE AND MILD/MODERATE COGNITIVE IMPAIRMENT



Results showed that the total number of children with cognitive impairment varied somewhat from one NV measure to another. While a slightly higher rate of impaired children was found for RPM (8/33), the same number of children performed below norms on Block Design and Matrix Reasoning (5/33). Almost all children displaying intellectual impairment on RPM showed a borderline cognitive impairment. Individual profiles of performance on each one of the four tests showed that the children with ASD displayed very heterogeneous performance on the four tests, with relative peaks and valleys of abilities depending on the NV measure. A detailed overview of individual performance is shown in Table 16.

TABLE 16. INDIVIDUAL PERFORMANCE OF CHILDREN WITH ASD ON NV MEASURES OF COGNITIVE ABILITIES

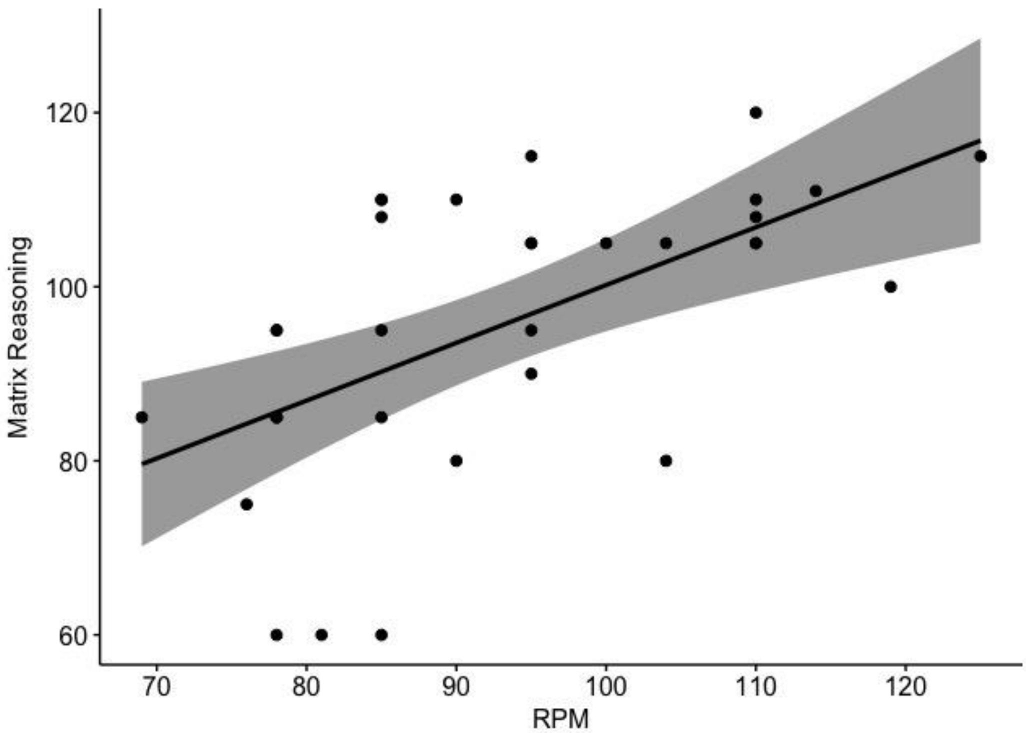
Code	RPM	Block Design	Matrix Reasoning	Picture Concepts
SCO	81	85	60	
RUG	95	105	105	100
ASC	100	105	105	75
VOR	110	103	108	
MTH	78	100	85	80
YLA	104	125	105	70
LWA	69	110	85	80
ODI	95	90	105	105
MON	110	110	120	110
CUT	110	80	105	100
CIP	78	65	85	
NUG	85	100	108	
NAF	85	85	110	
NOS	114	87	111	
ATE	78	70	95	80
EDT	95	120	95	
EPI	85	120	110	95
EMP	95	95	115	105
TUC	110	105	105	
GOT	85	115	95	85
ELO	85	100	85	85
AVI	119	110	100	115
MUG	110	110	110	115
JOS	78	95	95	55
FIZ	76	65	75	65
GHO	85	95	60	85
ADO	95	79	79	75
LPG	78	70	85	80
MIR	90	110	80	
LAT	104	100	80	90
MOI	90	110	110	105
YAT	78	55	60	55
LEC	125	92	115	

Note: Red boxes indicate mild / moderate impairment at ≤ 69 standard score, yellow boxes indicate borderline performance between 70 and 79 standard score

With the exception of two children (FIZ and YAT), who performed below norms on all four tests, all the other children showed either a relative valley of performance on one test (RPM, Matrix Reasoning or Picture Concepts) or on two tests (RPM and Block Design, RPM and Picture Concepts or Block Design and Picture Concepts).

Spearman correlation analyses were run between the four NVIQ measures (RPM, Block Design, Matrix Reasoning and Picture Concepts). Due to the high variability in individual performance showed in Table 16 no significant correlation was found between Block Design and RPM (r_s = .276, p = .120) or between Block Design and Matrix Reasoning (r_s = .325, p = .070). The correlation between Matrix Reasoning and RPM revealed a certain tendency toward significance (r_s = .603, p = .04), but as shown in Figure 14, the distribution was quite heterogeneous.

FIGURE 14. COMPARISON BETWEEN STANDARD SCORES ON MATRIX REASONING AND ON RPM (SPEARMAN CORRELATION WITH REGRESSION LINE)



Concerning Picture concepts we could run the analysis only on 23 children. Significant correlations were found with Block Design (r_s = .441, p = .035), RPM (r_s = .651, p = .001) and Matrix Reasoning (r_s = .627, p = .001).

From these results we can conclude that RPM, Block Design, and Matrix Reasoning may involve different abilities. Further evidence for this was the fact that, except for one child who had the exact same performance on the three NVIQ tasks (MUG), all children displayed at least a 10 standard score difference between one measure (which varied from child to child) and the others. More cautious conclusions should be drawn for the Picture Concepts subtest, since 1/3 of the children with ASD were unable to complete the task.

Before moving to conclusions and choosing which NV measures should be used for further analysis, we move to the last question in order to verify whether the performance of our ASD group on the four NV measures was related to language abilities.

5.3.2.3. NV measures and language abilities

Question 1f: *To what extent are RPM and the subtests of the PRI (Block Design, Matrix Reasoning, Picture Concepts) related to language abilities?*

Both RPM and the PRI are classified as NV measures of cognitive abilities. However, to our knowledge, no study has directly investigated the relation between performance on RPM, the three subtests of the PRI and linguistic abilities in children with ASD. A few studies in the literature have explored the relation between verbal abilities and NV tasks in the TD population. Concerning RPM, Fox & Charness, (2010) demonstrated that verbal abilities involved in performance on RPM which includes thinking out loud may increase performance by eleven points in older typically developing adults. Nader et al. (2016) advocated the possibility that subvocal reasoning may be involved in the resolution of RPM and that results similar to those obtained for TD adults could be obtained in high-functioning children with ASD. However, no study has investigated this hypothesis.

Concerning the three subtests of the PRI, it has been suggested that verbal mediation may occur during completion of the Picture Concepts task (since children are asked to verbalize their ‘self-explanations’), in contrast to Block Design and Matrix Reasoning (Houskeeper 2011; Sattler & Dumont, 2008). Block Design essentially involves visuospatial reasoning and spatial processing skills while the Matrix Reasoning subtest may be impacted by visual abilities more than by verbal abilities. If verbal reasoning is mobilized on a measure intended to be nonverbal, this may lead to inaccurate interpretation of test results, particularly in cases of language impairment. In particular for the Picture Concepts subtest, the scores may be correlated with performance representing verbal abilities. Matrix Reasoning has been described ostensibly as a measure of fluid intelligence (Williams et al., 2003) and although verbal mediation may occur during the process of pattern recognition and inductive reasoning, Houskeepeer (2011) has shown that verbal reasoning is only remotely involved in the performance of this task because examinees generally rely more consistently on visuospatial abilities and fluid reasoning processing to solve the items.

Considering that some studies have found a strong relation between NV tasks and linguistic abilities in TD subjects, it is therefore important to determine exactly which skills children with ASD are relying on when performing on RPM, Matrix Reasoning, Picture Concepts and Block Design so that the tests are utilized accurately and in a manner which can contribute to the interpretation of cognitive abilities in the ASD population. In the next section we will investigate possible relations between linguistic abilities and NV tasks in our group of ASD children. This analysis stems from our desire to use cognitive scores which are as nonverbal as possible, just as we have sought to use language measures capable of measuring as faithfully as possible structural language abilities.

Spearman correlation tests were run between the four NVIQ measures (RPM, Block Design, Matrix Reasoning and Picture Concepts) and measures of linguistic abilities on the performance of the ASD group (a composite score for vocabulary: VocR and VocP; a composite score for morphosyntax: MorsynR and MorsynP; the Phono score, and SR and NWR scores). Results are reported in Table 17.

TABLE 17. SPEARMAN'S CORRELATIONS BETWEEN NVIQ MEASURES AND TASKS EVALUATING STRUCTURAL LANGUAGE ABILITIES

NV measures		Language measures				
		Voc score	Morsyn score	Phono	NWR	SR
RPM	r_s	.487	.614	.421	.409	.469
	p	.004	< .001*	.021	.018	.006
	N	33	33	32	33	33
Block Design	r_s	.023	.141	.070	-.106	-.097
	p	.902	.457	.727	.578	.612
	N	33	33	32	33	33
Matrix Reasoning	r_s	.139	.261	.198	.088	.292
	p	.465	.164	.322	.643	.117
	N	33	33	32	33	33
Picture Concepts	r_s	.608	.615	.313	.263	.535
	p	< .001*	.001*	.076	.139	.001*
	N	23	23	22	23	23

*$p < .003$ with post-hoc Bonferroni correction for multiple comparisons

Results generally confirmed the expectations based on the literature: Block Design and Matrix Reasoning scores were not correlated with any measure of language abilities. Picture Concepts scores were significantly correlated with the vocabulary composite score, the morphosyntactic composite score and SR, indicating that verbal abilities are strongly involved in the performance of our group of children with ASD on this task (in line with Houskeeper, 2011). RPM scores on the other hand were significantly correlated only with morphosyntactic composite scores, although a tendency to significant correlation was found also with performance on vocabulary and SR.

5.3.3 Conclusions on cognitive abilities

The purpose of the present section was to verify the assumption that NV cognitive tasks, RPM and the PRI of WISC-IV, are the most adapted tools for evaluating cognitive abilities in relation to linguistic abilities in children with ASD, in contrast to both FSIQ and the other indices of WISC-IV. Our results confirmed this hypothesis. The group of 33 children with ASD displayed enhanced performance on NV measures, RPM and the PRI, in relation to the other measures of cognitive abilities. The number of children displaying a borderline + mild/moderate impairment was much higher when assessment was based on general measures of cognition, such as FSIQ of WISC-IV, or indices of cognitive abilities, such as VCI, PSI and WMI, than on NV measures. Moreover NV tools were essentially independent from language abilities. Based on these results we will favour the evaluation of cognitive abilities in children with ASD via the use of NV measures (RPM and PRI) rather than FSIQ, VCI, PSI and WMI.

Focusing on NV measures, we verified that the four tasks we selected evaluated different NV abilities, in order to avoid redundant effects in further analyses. We compared scores of our ASD group on RPM and the three subtests of the PRI. Our results showed that no significant correlations were found between the measures of RPM, Block Design and Matrix Reasoning. Individual performance showed high variability, with most children displaying at least some differences between two or more of these scores. This result motivated retaining RPM, Block Design and Matrix Reasoning in further analyses, since the observed discrepant performance was interpreted as the three tests involving different abilities.

Finally, to ensure a reliable evaluation of the relation between linguistic and cognitive abilities, we verified that the four NV measures were not (or at least only remotely) linked to verbal abilities. The score on Picture Concepts was the only one that showed strong correlations with more than one measure of language abilities, indicating, in line with what Houskeeper (2011) demonstrated for TD children, that children with ASD strongly rely on verbal skills to perform this task as well. While performance on Block

Design and Matrix Reasoning did not display any relation with language abilities, the RPM results correlated with the morphosyntactic scores, indicating (in line with Nader et al., 2016) that this test may involve some level of verbal subvocalisation. However, while for Picture Concepts verbalisation is an integral part of the task, for RPM we can only suppose that (some) children with ASD (sometimes) subvocalize their reasoning process. In conclusion, for further analysis of structural language and cognitive abilities profiles of children with ASD, we decided to keep three NV measures (Block Design and Matrix Reasoning and RPM) and to exclude the Picture Concepts subtask. This decision stems from the high number of children who could not complete the latter task and, more importantly, from our desire to use cognitive scores that are as nonverbal as possible. Moreover, this choice seems to go in the same direction as the new version of the WISC (WISC-V), which has excluded Picture Concepts (previously included in the PRI) from both the obligatory tasks of *Fluid Reasoning Index (FRI)* of the new primary index scales and from the new *Nonverbal Index* (NVI) of the ancillary index scales.

5.3.4 Children assessed via the EDEI-R

Four children in the ASD group were assessed via the EDEI-R psychometric battery. Table 18 reports individual scores on RPM and FSIQ, and verbal and nonverbal indices of EDEI-R.

TABLE 18. INDIVIDUAL PERFORMANCE OF CHILDREN WITH ASD ON RPM AND EDEI-R

Child	RPM	FSIQ of EDEI-R	V index EDEI-R	NV index EDEI-R
SIM	114	89	78	104
KEV	78	105	67	82
EVA	95	67	59	80
ROS	90	74	54	95

Note: Red boxes indicate mild / moderate impairment at ≤ 69 standard score, yellow boxes indicate borderline performance between 70 and 79 standard score

Similarly to what we found for WISC-IV scores, all four children displayed enhanced NV abilities in comparison to Verbal index and FSIQ scores. Only one child showed borderline performance on RPM, while on the NV index of EDEI-R all four children performed in the norms. Due to the very small number of children, we could not run any further statistical analysis. These four children will also be treated separately in the next chapter, which will concentrate on the description of structural language and NV ability profiles in children with ASD.

5.4 Descriptive classification of cognitive profiles

Here we report the descriptive classification of our group of ASD children on the basis of the NV measures selected, RPM, Block Design and Matrix Reasoning, for the 33 children assessed via the WISC-IV and RPM, and the NV index for the four children assessed via EDEI-R. We considered children as having an intellectual impairment when they peformed below norms on at least two (out of three) NV tasks. With this criterion seven children would be considered as intellectually impaired among our ASD group: CIP, ATE, JOS, FIZ, LPG, ADO, YAT. If we applied descriptive classification of IQ, we notice that no child, despite showing different performance on the three measures, displayed a discrepancy of more than two consecutive steps of the IQ classification between the three scores. We considered children having relative peak/valley of performance on a task when they performed with a difference of two consecutive categories of the IQ classification (e.g. SCO showed a relative valley of performance on Matrix Reasoning; YAT displayed a relative peak of performance on Matrix Reasoning). To facilitate interpretation, we copy here the descriptive table from Chapter 2 (Table 5).

<table>
<tr><th>Classification</th><th>Standard scores</th><th></th></tr>
<tr><td>Very superior</td><td>≥ 130</td><td></td></tr>
<tr><td>Superior</td><td>120-129</td><td></td></tr>
<tr><td>High average</td><td>110-119</td><td></td></tr>
<tr><td>Average</td><td>90-109</td><td></td></tr>
<tr><td>Low average</td><td>80-89</td><td></td></tr>
<tr><td>Borderline</td><td>70-79</td><td></td></tr>
<tr><td>Mild ID</td><td>55-69</td><td rowspan="4">Extremely Low IQ</td></tr>
<tr><td>Moderate ID</td><td>40-54</td></tr>
<tr><td>Severe ID</td><td>25-39</td></tr>
<tr><td>Profound ID</td><td>< 25</td></tr>
</table>

5.5 Language abilities and NV abilities in relation to developmental factors

Before moving to the general conclusions of this chapter we will briefly investigate the possible relation between developmental factors (age of first word and age of first sentence), structural language measures (NWR and SR) and NV cognitive measures (RPM, Block Design and Matrix Reasoning). Results are reported in Table 19.

TABLE 19. SPEARMAN'S CORRELATION BETWEEN SR, NWR, RPM, BLOCK DESIGN, MATRIX REASONING AND DEVELOPMENTAL FACTORS

	Age of 1st word (months)	Age of 1st sentence (months)
SR	r_s = - .205, p = .269	r_s = - .631, p = .002*
NWR	r_s = - .229, p = .216	r_s = - .511, p = .015
RPM	r_s = - .282, p = .124	r_s = - .452, p = .034
Block Design	r_s = .014, p = .942	r_s = - .247, p = .295
Matrix Reasoning	r_s = - .145, p = .462	r_s = .052, p = .824

*p <.010 with post-hoc Bonferroni correction for multiple comparisons

Age of 1st word was not correlated with either measures of structural language abilities or measures of NV cognitive abilities. On the other hand, age of 1st sentence was significantly correlated with both SR and RPM. These results are only partially in line with previous findings in the literature. Age of first word was not a strong predictor for either language or cognitive abilities in our group of children with ASD (contra Kover et al., 2016; Mayo et al. 2013). For age of first sentence no significant correlation was found either with Block Design or with Matrix Reasoning scores (contra Wodka et al., 2013). However, as anticipated in Chapter 4, since data regarding age of first sentence could be retrieved for only 23/37 children, group results could have been biased by the low number of children. For this reason we looked at individual performance. Taking the threshold for late onset of first sentence production at > 30 months (Kenworthy et al., 2012), we found that 5/23 children (22% of the sample) displayed normal performance on SR despite their first sentence having been produced considerably after 30 months of age (48 – 60 months). Individual results thus indicated that children with ASD may develop normal structural language abilities despite late structural language onset.

5.6 General conclusions and discussion

The present chapter demonstrated that repetition tasks (SR and NWR) and NVIQ measures (RPM, Matrix Reasoning and Block Design) were likely to be the most domain-specific and the most appropriate tools for measuring structural language and nonverbal abilities in children with ASD. On the one hand standardized tests were so densely intertwined with multiple factors that they did not enable detailed analyses of grammatical abilities, leading to an underestimation of structural language abilities of some of our children with ASD (16% for morphosyntactic abilities and 26% for phonological abilities, which is not negligible). These results confirm the hypothesis of previous studies arguing for evaluation of linguistic abilities in ASD via the use of specific tools developed for minimizing working memory effects (NWR and SR) and reducing the effect of pragmatic impairment and use of language in a conversational context (SR) (see Prévost et al., 2018; Tuller et al., 2017; Wittke et al., 2017). On the other hand global measures of cognitive abilities (FSIQ scores) and domain-specific cognitive measures (verbal index, working memory index, etc.) did not prove to be well-suited to capture the unique cognitive phenotype of children with ASD. Our results confirmed previous results (Courchense et al., 2015; Barbeau et al., 2013; Nader et al., 2016; a.o) suggesting that tasks evaluating NV abilities (fluid reasoning and visuospatial abilities) may actually highlight the “real” capacities of children with ASD. For our analysis we retained RPM, Block Design and Matrix Reasoning since they did not correlate with each other and they did not show significant correlations with language abilities. This was not the case of the Picture Concepts subtest. These results may cast new light on the debate over what aspects of autistics' intelligence these tests do actually measure. Future studies might, in particular, investigate the performance of children with autism on Matrix Reasoning and RPM to see if the heterogeneity of individual results found in our population sample can be replicated on a larger sample. Further analysis is also needed on the relation between RPM and language abilities on a larger group of children, including a direct evaluation of the possible influence of subvocalic language on RPM performance. Although there was no significant correlation between RPM scores and language abilities, some tendencies towards significance were found. In this vein it is interesting to report that even in the Raven’s manual (Raven, 1998) the authors discussed the possibility that individuals may use subvocalisation in performing RPM. No study, however, did prove this hypothesis. We think that subvocalisation can be used to perform this task, but that its use is not compulsory; children can also focus on spatial/fluid reasoning abilities. A possible evidence of this is that children with SLI perform in the norm on RPM although they display language impairment.

Chapter VI
Structural language and NV cognitive ability profiles in children with ASD

6.1 Introduction

The aim of the present chapter was to put the heterogeneity of ASD to the forefront by investigating **whether clear profiles related to structural language and NVIQ abilities emerge when we apply the measures derived from the answer to (1) to a population sample taken from the entire spectrum.** In order to explore this question, we have, in fact, explicitly argued for the use of specific measures of formal language abilities and of NV cognitive abilities (Chapter 5). In this chapter we will consider possible links between these measures and factors previously argued to predict these abilities (age of first word and age of first sentence). Then, we will explore structural language / nonverbal ability profiles by using an integrative approach that takes in consideration both linguistic and cognitive abilities.

6.2 Methods

6.2.1 Participants

The results presented in the present chapter included only the thirty-three children with ASD who were assessed through the WISC-IV task (age M = 9.08, SD = 18.7) (see Chapter V). According to the ICD-10 criteria, the group included 13 children with Autistic disorder, 12 children with PDD-NOS, and 8 children with Asperger's Syndrome. The group was composed of 1 girl and 32 boys. The other four children were not included because their nonverbal cognitive abilities were assessed via the EDEI-R and this test does not provide the same information as the WISC-IV (notably, the EDEI-R NV score does not entail any Block Design or Matrix Reasoning type task). The results of these four children will be analysed a posteriori by matching their profiles to the ones detected by our analysis of structural language/NV abilities on the other 33 children.

6.2.2 Materials and procedure

Following the results of Chapter 5, for the present analysis structural language abilities will be ascertained exclusively on the basis of scores from the two experimental repetition tasks targeting specific aspects of structural language (SR and NWR), and cognitive abilities will be based on RPM and the two subtests of the PRI, Block Design and Matrix Reasoning. For the purposes of this study, the ADOS severity score, the CARS and the ECA-R global score were used as measures of autism severity.

6.2.3 Data analysis

The main aim of our analysis was to describe structural language/nonverbal ability profiles in our population sample. In order to attain this goal the analysis was divided in three parts. Due to the limited number of participants a Principle Component Analysis (PCA) was first conducted on the totality of the factors that may contribute in describing the variability of our dataset. This allowed us to detect which variables accounted the most for the data variability and to reduce the number of parameters that could create confounding effects. Following this analysis, severity of autism scores were excluded.

Two types of cluster analyses were then conducted on the data selected from the preceding analysis: the two clustering methods proposed an integrated approach, based on both linguistic and cognitive abilities in describing profiles in children with ASD. The first type of cluster analysis was run through a *hard clustering method (partitioning clustering)*: in hard clustering each data point either belongs to a cluster completely or not. The second type of cluster analysis was run through a *soft clustering method (model-based clustering)*: in soft clustering instead of putting each data point into a separate cluster, a probability or likelihood of that data point to be in its cluster is assigned.

6.3 Results

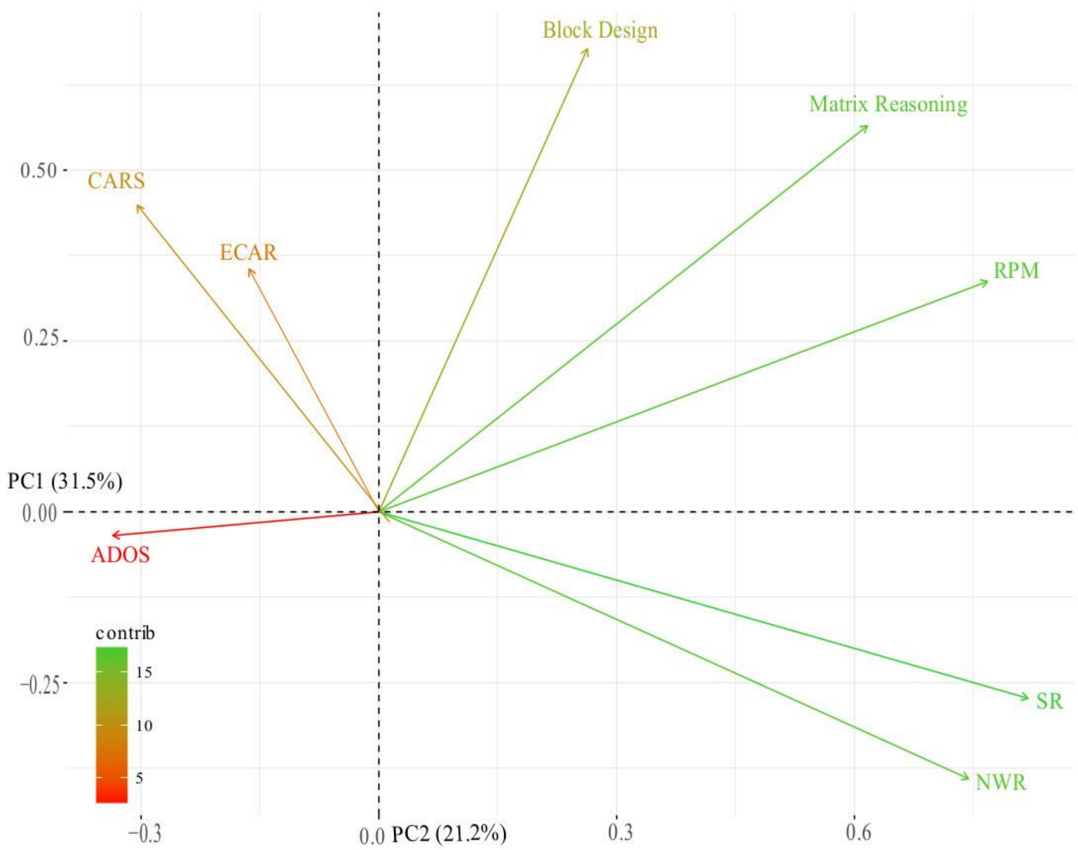
6.3.1 Reducing the number of factors in describing language and nonverbal ability profiles

Question 1*: To what degree do extra-linguistic and extra-cognitive factors, notably severity of autism symptoms, have an effect on structural language abilities and NV intelligence?*

In order to properly investigate the relationships between language and nonverbal abilities and to alleviate confounding effects of intercorrelations among the different parameters, it was useful to project the data along orthogonal principal components (PCs). Besides measures of structural language (SR and NWR) and nonverbal abilities (RPM, Block Design and Matrix Reasoning), the PCA analysis included scores for severity of autism (ADOS severity scores, CARS and ECA-R). The data were scaled for normalization purposes due to the presence of different kinds of measures, using a Bayesian linear regression method (Buuren & Groothuis-Oudshoorn, 2010).

The PCA (Figure 15) revealed two principal components (PCs), which together accounted for 52.7% of the data variability. Table 20 describes the variables' contribution to each PC, highlighting the most significantly associated variables with a given principal component. Due to multiple comparisons, post-hoc *t*-tests included Bonferroni corrections in order to avoid type 1 error; results were considered significant at $p < .006$.

FIGURE 15. THE PRINCIPLE COMPONENT ANALYSIS (PCA) INCLUDING ALL THE PARAMETERS



Note: The contribution bar highlights the most important variables that contribute to explaining the variability in the data set

TABLE 20. VARIABLES CONTRIBUTING TO PCS – CORRELATION AND *P*-VALUES FOR PC1 AND PC2

PC1 (31.5%)	**Correlation**	***p*-value**
SR	.819	< .001
RPM	.767	< .001
NWR	.744	< .001
Matrix Reasoning	.616	< .001
PC2 (21.2%)	**Correlation**	***p*-value**
Block Design	.677	< .001
Matrix Reasoning	.564	< .001

Note: Significance threshold after post-hoc correction – Bonferroni was $p \leq .006$

Results suggested that a combination of both structural language and NVIQ measures accounted for the identification of PC1 (31.5%). All measures of structural language and NV abilities contributed to the first dimension, except for Block Design. PC2 (21.2%) relied exclusively on nonverbal measures (Block Design and Matrix Reasoning). Severity of autism scores did not contribute significantly to either of the PCs: PC1/ADOS (r = -.335, p = .056); PC1/ECAR (r = -.217, p = .224); PC1/CARS (r = -.304, p = .085); PC2/ADOS (r = -.034, p = .847); PC2/ECAR (r = .335; p = .056); PC2/CARS (r = .478; p = .010).

6.3.2 Interim discussion

The results on the PCA allowed us to reduce the number of components necessary to define structural language / nonverbal profiles in children with ASD. In particular, autism severity scores could be excluded from further analyses due to the fact that they were only marginally involved in explaining the variability of the dataset. The PCA analysis showed that the measures most highly involved in explaining our dataset were a combination of both nonverbal IQ and structural language abilities.

6.3.3 Cluster analyses

Due to the great heterogeneity of results on standardized tasks reported in Chapter 5 and the labelling issue that was raised earlier in Chapter 1 for the identification of structural language profiles and in Chapter 2 for the identification of cognitive profiles, we decided to unsupervised exploratory analysis based on data mining that tries to identify structures within the data. Cluster analysis is a statistical classification technique in which a set of objects or points with similar characteristics are grouped together in clusters. It encompasses a number of different algorithms and methods that are all used for grouping objects of similar kinds into respective categories. The aim of cluster analysis is to organize observed data into meaningful structures in order to gain further insight from them. Cluster analysis can be considered a tool for exploratory data analysis that is aimed at sorting different objects into meaningful groups in such a way that the degree by which these objects are associated is at the maximum if they belong to the same group and at the minimum if they do not. Cluster analysis is used to bring to light the hidden structures or relationships within a data set, independent of the need to explain or interpret what this relationship is. In essence, cluster analysis is used only to discover the structures found in data; it does not explain why these structures or relationships exist. Such explanations must be described and motived by the researcher.

In the present study, we conducted multiple cluster analyses in order to more fully identify and describe structural language/nonverbal ability profiles in children with ASD. First, we wanted to verify if cognitive profiles would distribute homogeneously on structural language abilities and if profiles of structural language abilities would distribute homogeneously on NV cognitive abilities. We ran two cluster analyses, on the basis of each type of ability. We then verified a posteriori how NV cognitive profiles and structural language ability profiles distributed over the results of the clusterings. For this analysis, K-means clustering method was adopted. This kind of clustering consists of unsupervised learning, which is used for unlabelled data (i.e., data without defined categories or groups). The goal of this algorithm is to find groups in the data, with the number of groups represented by the variable K. The algorithm works iteratively to assign each data point to one of K groups based on the features that are provided. Data points are clustered based on feature similarity and they distribute around a centroid (the point that is geometrically the centre of a cluster). Each centroid of a cluster is a collection of feature values, which define the resulting groups. Examining the centroid feature weights can be used to qualitatively interpret what kind of group each cluster represents. A function, independent of the K-means algorithm, determines the optimal number of clusters using different methods, within cluster sums of squares, average silhouette and gap statistics (Kassambara, 2017). Data were already normalized from the previous PCA analysis. For each analysis, labels indicate the child codes and the relative diagnostic classification (ICD-10) in parenthesis: A = Autistic disorder, P = PDD-NOS, AS = Asperger’s Syndrome.

Figure 16 shows the results of the cluster analysis for structural language and Figure 17 shows results of the cluster analysis for NVIQ abilities. The optimal number of clusters was automatically estimated to be *two* in both analyses. For each cluster analysis the algorithm returned the characteristics of each centroid (two for the cluster analysis in Figure 16 and two for the cluster analysis in Figure 17) which allowed us to describe and define the groups. In the first cluster analysis (language abilities), children were clearly separated into two groups of the basis of their scores on NWR and SR. The centroid's scores of the cluster in the upper right corner were NWR = 0.49 and SR = 0.58, while for the cluster in the bottom left corner they were NWR = -1.28 and SR = -1.34. In the second cluster analysis (NV abilities) children were separated into two groups of the basis of their scores on RPM, Block Design and Matrix Reasoning. The scores of the centroid in the right cluster were: RPM = 0.56, Block Design = 0.42 and Matrix Reasoning = 0.62, while for the left cluster they were: RPM = -0.86, Block Design = -0.65 and Matrix Reasoning = -0.95. After data normalisation the threshold for a score in the norms was fixed at 0. The centroids' scores showed that a performance above the threshold was displayed by the cluster in the upper right corner in Figure 16 for language abilities and by the right cluster in Figure 17 for NV abilities. The other two clusters in Figures 16 and 17 displayed a performance below the threshold.

FIGURE 16. K-MEANS CLUSTER ANALYSIS ON THE MEASURES OF SR AND NWR

(nb of clusters estimated by optimal solution = 2, distance matrix = Euclidean, ellipse type = euclid)

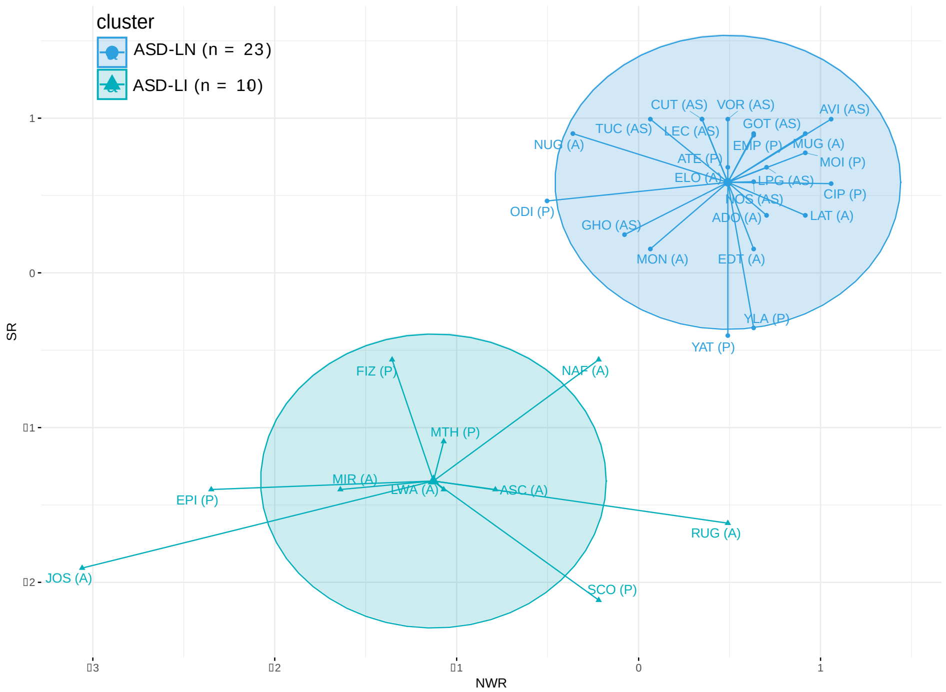
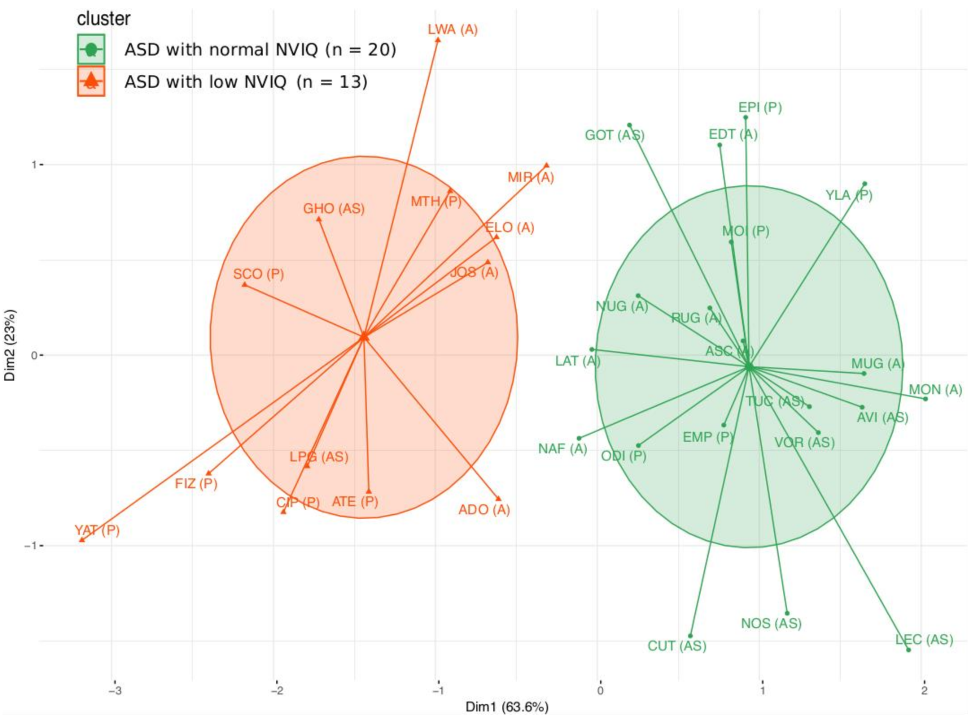


FIGURE 17. K-MEANS CLUSTER ANALYSIS ON THE MEASURES OF NVIQ

(nb of clusters estimated by optimal solution = 2, distance matrix = Euclidean, ellipse type = euclid)



The description of group performance (mean and SD) for each cluster appears in Table 21. Together with the identification of the centroids rescaled to original data, these results allowed us to better describe and justify the labels for the two clusters of language abilities as ASD-LI and ASD-LN and for the two clusters of NVIQ abilities as low NVIQ and normal NVIQ.

TABLE 21. CHARACTERISTICS OF THE TWO CLUSTERS OF LANGUAGE ABILITIES AND OF THE TWO CLUSTERS OF COGNITIVE ABILITIES, MEAN AND (SD)

Language Clusters	**SR (%)** [a]	**NWR (%)** [b]	
ASD-LN (*n* = 23)	86 (13)	90 (5)	
ASD-LI (*n* = 10)	24 (16)	67 (14)	
Cognitive Clusters	**RPM (standard score)** [c]	**Block Design (standard score)** [c]	**Matrix Reasoning (standard score)** [c]
Normal NVIQ	99 (13)	103 (12)	104 (9)
Low NVIQ	78 (5)	80 (17)	79 (13)

[a] Cut-off for low performance on SR was established at < 78 % of correct repetitions (Tuller et al., 2018)
[b] Cut-off for low performance on NWR was established at < 77 % of correct repetitions (Tuller et al., 2018)
[c] Cut-off for low NVIQ was established at < 80 standard score

When we looked at the distribution of NV profiles in the language ability clusters and at the distribution of language profiles in the NV ability clusters, we noticed that both children with normal and low NVIQ were present in each of the two language ability clusters and that both children with normal and low language abilities were present in each of the two NV ability clusters. Table 22 describes the NV cognitive profiles of the two clusters of children divided on the basis of their language abilities and Table 23 describes the linguistic profiles of the two clusters of children divided on the basis of their NV abilities.

TABLE 22. CHARACTERISTICS OF THE TWO LANGUAGE ABILITY CLUSTERS FOR NVIQ MEASURES, MEAN (SD) AND MINIMUM – MAXIMUM VALUES

	ASD-LI **(*n* = 10)**	**ASD-LN** **(*n* = 23)**
RPM [a]	83 (9.3) 69 - 100	97 (14.1) 78 - 125
Block Design [a]	98 (16) 65 - 120	95 (18) 55 - 125
Matrix Reasoning [a]	91 (16.7) 60 - 110	97 (16) 60 - 120

[a] Cut-off for low NVIQ was established at < 80 standard score

TABLE 23. CHARACTERISTICS OF THE TWO NV ABILITY CLUSTERS FOR STRUCTURAL LANGUAGE MEASURES, MEAN (SD) AND MINIMUM – MAXIMUM VALUES

	Low NVIQ **(*n* = 13)**	**Normal NVIQ** **(*n* = 20)**
SR [a]	56 (33) 0 - 90	74 (30) 16 - 100
NWR [b]	79 (17) 40 - 98	85 (11) 50 - 98

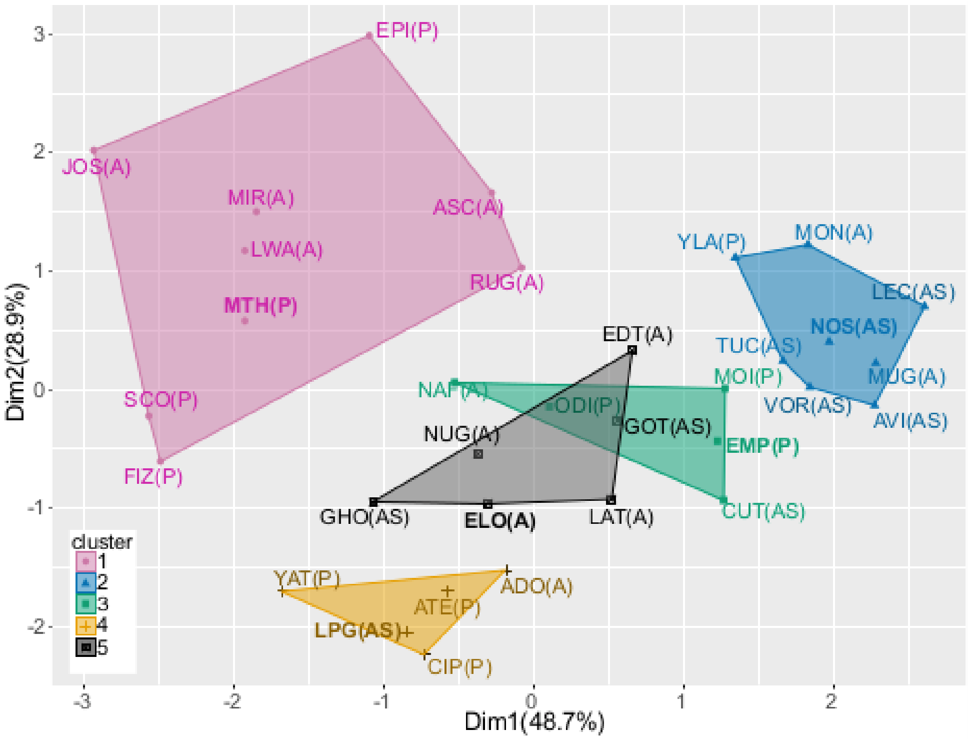
[a] Cut-off for low performance on SR was established at < 78 % of correct repetitions (Tuller et al., 2018)
[b] Cut-off for low performance on NWR was established at < 77 % of correct repetitions (Tuller et al., 2018)

In order to investigate whether clear profiles would emerge from the intersection of language abilities and NV cognitive abilities, we used an integrative approach that takes into consideration both abilities. For the first analysis, the Partitioning Around Medoids (PAM) non-hierarchical k-medoid clustering method was adopted. This algorithm is more adapted for small datasets and it can manipulate noisy data and outliers (for a detailed description of the PAM clustering method, see Kaufman & Rousseeuw, 1990). By empirically taking one of the objects of the cluster as the prototype (medoid), the k-medoid algorithm allows for the identification and description of the main features of the other objects included in the same cluster. The term "medoid" refers to an object within a cluster for which average dissimilarity between it and all the other members of the cluster is minimal. These objects (one per cluster) can be considered as a representative example of the members of that cluster (Soni & Patel 2017).

The clustering variables for nonverbal measures and language measures motivated above by the PCA analyses were RPM, Block Design, Matrix Reasoning, SR and NWR. The number of clusters was automatically estimated by optimum average silhouette width (Hennig & Liao, 2013). Convex frame type representation of the clusters was chosen to outline all the objects that showed profiles similar to the prototype (medoid). Finally, a Silhouette (*Si*) cluster validation approach was used to evaluate the goodness of clustering algorithm (Brock et al., 2011). The average *Si* coefficient width measured how well an observation was clustered and it estimated the average distance between clusters. Values near 1 indicated that the observations were very well clustered; values near 0 meant that the observations were situated between two clusters.

Figure 18 shows the results of the cluster analysis for both structural language and NVIQ abilities. The optimal number of clusters was automatically estimated to be *five* (nb of clusters estimated by optimal solution = 5, distance matrix = Manhattan, ellipse type = convex). Each cluster medoid appears in bold face.

FIGURE 18. PARTITIONING AROUND MEDOIDS (PAM) CLUSTER ANALYSIS ON THE MEASURES OF SR, NWR, BLOCK DESIGN, MATRIX REASONING AND RPM



The description of group performance (median and min/max values) for each cluster and the identification of the medoids allowed us to identify and tentatively name the five clusters as ASD-LI with low average / average NVIQ abilities (cluster 1), ASD-LN with high average / average NVIQ abilities (cluster 2), ASD-LN with average NVIQ abilities (cluster 3), ASD-LN with low / low average NVIQ abilities (cluster 4) and ASD-LN with low average / average NVIQ abilities (cluster 5). The analysis revealed two clear structural language/nonverbal ability profiles (clusters 2 and 4) and three less clear structural language/nonverbal ability profiles (clusters 1, 3 and 5). Table 24 describes the performance for each cluster.

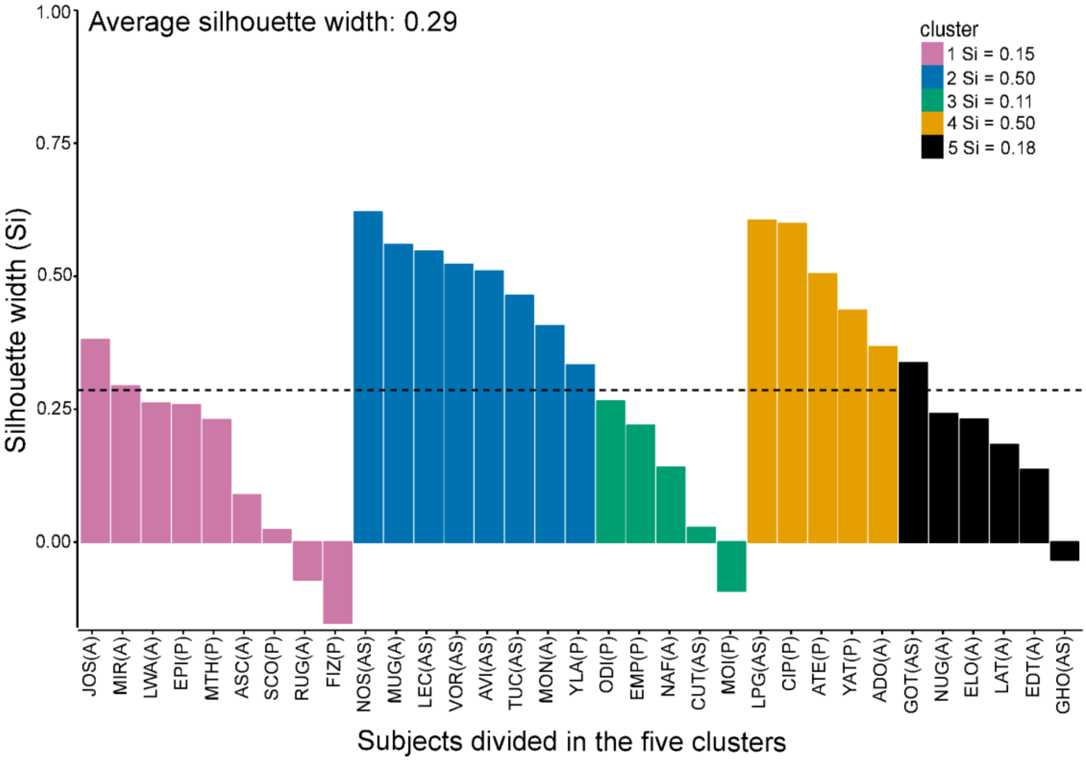
TABLE 24. THE MAIN CHARACTERISTICS OF THE 5 CLUSTERS (MEDIAN, MIN AND MAX VALUES)

	Cluster 1 *n* = 9	Cluster 2 *n* = 8	Cluster 3 *n* = 5	Cluster 4 *n* = 5	Cluster 5 *n* = 6
NWR (percentage) [a]	68 40-90	91 84-98	88 76-96	92 90-98	91 78-96
SR (percentage) [a]	23 0-50	98 56-100	93 50-100	86 55-90	93 73-97
RPM (standard score) [b]	81 68-100	110 104-124	95 85-110	78 78-95	85 85-105
Block Design (standard score) [b]	105 62-120	107 105-125	90 80-110	69 52-80	100 95-120
Matrix Reasoning (standard score) [b]	85 60-110	109 100-124	110 105-115	85 60-95	90 60-110
Structural language / NVIQ ability profiles	ASD-LI with low average/ average NVIQ abilities	ASD-LN with average / high average NVIQ abilities	ASD-LN with average NVIQ abilities	ASD-LN with low / low average NVIQ abilities	ASD-LN with low average / average NVIQ abilities

[a] The repetition cut-off rate was established at 77 % for NWR and 78 % for SR, which corresponds to high levels of diagnostic accuracy for language impairment (Tuller et al. 2018).
[b] The cut-off for RPM, Block Design and Matrix Reasoning was < 80 standard scores.

As shown by the location of the points in the convex ellipses, the five clusters involved a certain level of variability. Some subjects were relatively near the medoid (the prototypical object of the cluster), while others seemed to be collocated fairly far away. This graphical realization of the heterogeneity of our group was confirmed by the results of the *Si* cluster validation, which showed a fairly low average silhouette coefficient width at 0.29 (Figure 19). The average silhouette coefficient width measures how well all observations are clustered and estimate the average distance between clusters. In the legend we reported the average silhouette width for each of the five clusters.

FIGURE 19. CLUSTER SILHOUETTE PLOT

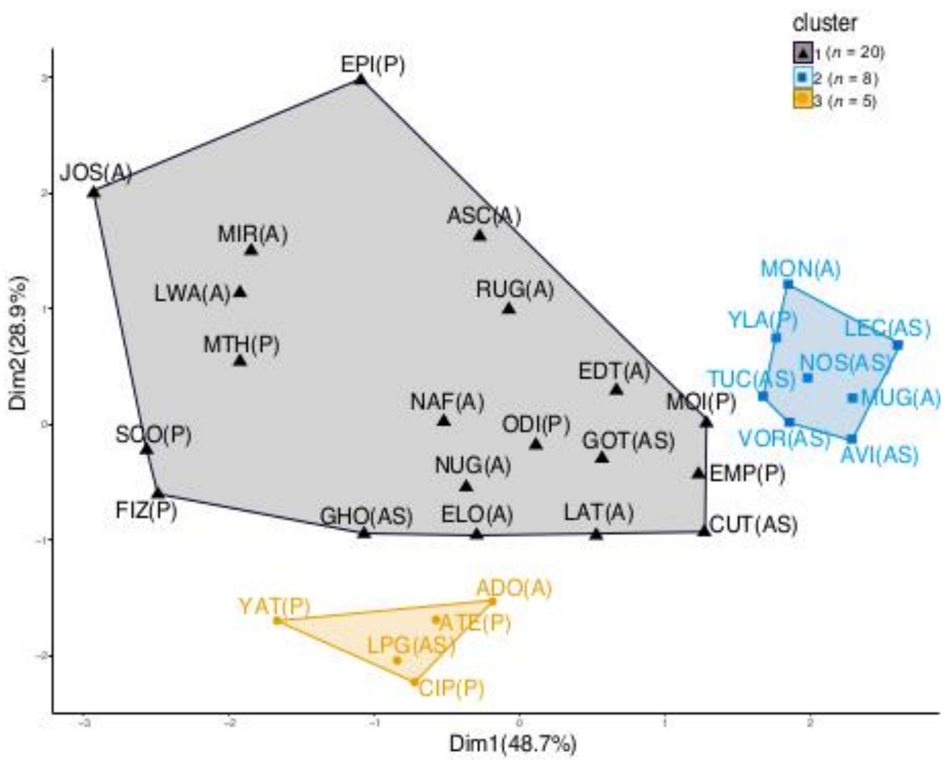


Note: The dotted line represents the average silhouette coefficient width (= 0.29)

However, not every cluster showed a low rate of confidence. Cluster 2 (ASD-LN with high average / average NVIQ abilities) and cluster 4 (ASD-LN with low / low average NVIQ abilities) both displayed an average width of 0.50, which indicates good clustering structure. On the other hand, cluster 1 (ASD-LI with low average / average NVIQ), cluster 3 (ASD-LN with average NVIQ), and cluster 5 (ASD-LN with low average / average NVIQ) were respectively at 0.15, 0.11 and 0.11, which indicates that no substantial structure could be found in any of these clusters. These three ill-defined clusters were fuzzy for the same reason: the profiles of the subjects involved were too heterogeneous to be well clustered together and although a general tendency was found in each cluster (medoid), this was not sufficient to be representative of the entire sample. As shown in Table 24, the minimum and maximum values for clusters 1, 3 and 5 ranged from very low scores to very high scores on every measure.

It could be that a hard clustering method such as the PAM algorithm "forced" each subject to end up in the cluster that displayed the fewest dissimilarities, even when they did not really share the exact same profile. In order to validate the results that emerged from the first cluster analysis, we ran a model-based cluster analysis on the same variables. Contrary to hard clustering methods, model-based clustering considers the data as coming from a distribution that is a mixture of two or more clusters. Unlike PAM, model-based clustering uses soft assignment, where each data point has a probability of belonging to each cluster (Scrucca et al., 2016). In this way only very strong profiles should emerge in the model as well-defined clusters, while more heterogeneous profiles should end up in a fuzzy group. The algorithm automatically estimates the model best suited for describing the dataset, the optimal number of clusters, and the probability for each observation to belong to each cluster. Results of the model-based cluster analysis are shown in Figure 20 (nb of clusters estimated by optimal solution = 3, clustering model selection = "VII", ellipse type = convex, cluster assignment for each observation = classification).

FIGURE 20. MODEL-BASED CLUSTER ANALYSIS ON THE MEASURES OF SR, NWR, BLOCK DESIGN, MATRIX REASONING AND RPM



The algorithm automatically selected a model with three clusters. Two strong clusters (2 and 3) clearly emerged. These clusters corresponded to the two well-defined clusters found in the previous PAM analysis, the ASD-LN with high average / average NVIQ abilities cluster and the ASD-LN with low / low average NVIQ abilities cluster. Cluster 1, which completed the analysis, was the result of the union of the 20 children from the three heterogeneous clusters previously detected with the PAM algorithm. Since soft constraints reflect the uncertainty associated with a priori knowledge about pairs of points that could or could not belong to the same cluster, we can argue that this second analysis validated our previous results.

But why is cluster 1 fuzzy? In Table 25, we report individual performance of the children included in this fuzzy cluster.

TABLE 25. INDIVIDUAL SCORES OF THE 20 CHILDREN COMPOSING THE FUZZY CLUSTER

Child Code	NWR (%)	SR (%)	RPM (standard score)	Block Design (standard score)	Matrix Reasoning (standard score)
SCO (P)	80	0	81	85	60
RUG (A)	90	16	95	105	105
ASC (A)	72	23	100	105	105
MTH (P)	68	33	78	100	85
LWA (A)	68	23	69	110	85
ODI (P)	76	83	95	90	105
CUT (AS)	88	100	110	80	105
NUG (A)	75	97	85	100	108
NAF (A)	80	50	85	85	110
EDT (A)	92	73	95	120	95
EPI (P)	50	23	85	120	110
EMP (P)	92	97	95	95	115
GOT (AS)	92	97	85	115	95
ELO (A)	90	87	85	100	85
JOS (A)	40	7	78	95	95
FIZ (P)	64	50	76	65	75
GHO (AS)	82	76	85	95	60
MIR (A)	60	23	90	110	80
LAT (A)	96	80	104	100	80
MOI (P)	96	93	90	110	110

Note: Scores below the thresholds are highlighted in grey

Focusing first on structural language, "fuzziness" came from two facts: (1) This cluster incorporated all of the children who displayed impaired structural language, phonology and/or morphosyntax (*n* = 14), as well as some children with unimpaired structural language (*n* = 6). (2) Of the 14 children with LI, five showed spared phonology and impaired morphosyntax, two spared morphosyntax and impaired phonology and 7 had both impaired morphosyntax and phonology. Focusing on NVIQ, the 6 children with normal language abilities displayed heterogeneous NVIQ, with at least 10 standard scores difference between one measure, which varied from child to child, and the others. Moreover, 6 of the 14 children with LI had heterogeneous NVIQ scores, with valleys of performance: two with impairment only in Matrix Reasoning (SCO and GHO) and three only in RPM (MTH, LWA and JOS). Recall that a relative peak/valley is an area in which an individual excels/fails compared with other areas of cognitive abilities in which (s)he performs worse/better (at least a difference of two steps in the IQ classification; see section 5.4.). Only one child, FIZ, could be identified as ASD-LI with low NVIQ.

Children in the two distinct clusters showed general homogeneity in their performance, both for NVIQ and for structural language, with three exceptions. These were children ("YLA" and "MON" for the ASD-LN with high average / average NVIQ cluster and "YAT" for the ASD-LN with low / low average NVIQ cluster) who displayed an impaired score in SR. The drop of performance of these three children for this task was due to a lack of collaboration and/or attention during the testing phase. YLA and MON were particularly distracted during the test, while YAT performed very badly in the second half of the test. Similarly to what we said before for the selective drop of performance on NVIQ measures in the fuzzy cluster, we will consider YLA and MON as LN. We cannot draw the same conclusions for YAT, whose performance was so singular that assigning him the label of LN might mean overestimating his language abilities. As independent evidence, if we look at the scores of these three children on target structure in SR, YLA and MON performed above the threshold, while YAT did not (see section 7.3.3.3 for further description of the Target structures score). For this reason we will consider YAT performance on SR as impaired. Table 26 displays individual results for children included in the two strong clusters.

TABLE 26. INDIVIDUAL SCORES OF THE 13 CHILDREN COMPOSING THE TWO STRONG CLUSTERS

Child Code	NWR (%)	SR (%)	RPM (standard score)	Block Design (standard score)	Matrix Reasoning (standard score)
ASD-LN with high average / average NVIQ					
VOR (AS)	90	100	110	103	108
YLA (P)	92	57	104	125	105
MON (A)	84	73	110	110	120
NOS (AS)	92	87	114	87	111
TUC (AS)	84	100	110	105	105
AVI (AS)	98	100	119	110	100
MUG (A)	96	97	110	110	110
LEC (AS)	88	100	125	92	115
ASD-LN with low / low average NVIQ					
CIP (P)	98	87	78	65	85
ATE (P)	90	90	78	70	95
ADO (A)	93	80	95	79	79
LPG (AS)	93	90	78	70	85
YAT (P)	90	55	78	55	60

Note: Scores below the thresholds are highlighted in grey

Since most discrepancies were found for language abilities (scores for SR and NWR did not go in the same direction), we ran two separated cluster analyses for morphosyntax and phonology only on the children belonging to the fuzzy cluster to verify whether clear profiles of abilities would emerge. In this case we used the K-means analysis, because we wanted to force our dataset to divide the fuzzy group into clusters of performance. The main aim was to focus on SR and NWR separately to see how the algorithm would separate the group. Figure 21 shows the results of the cluster analysis for SR and NVIQ abilities and Figure 22 shows the results of cluster analysis for NWR and NVIQ. The optimal number of clusters was automatically estimated to be *two* for each analysis (nb of clusters estimated by optimal solution = 2, distance matrix = Manhattan, ellipse type = confidence; for both analyses).

FIGURE 21. CLUSTER ANALYSIS (K-MEANS) ON THE MEASURES OF SR, BLOCK DESIGN, MATRIX REASONING AND RPM ONLY ON THE FUZZY CLUSTER

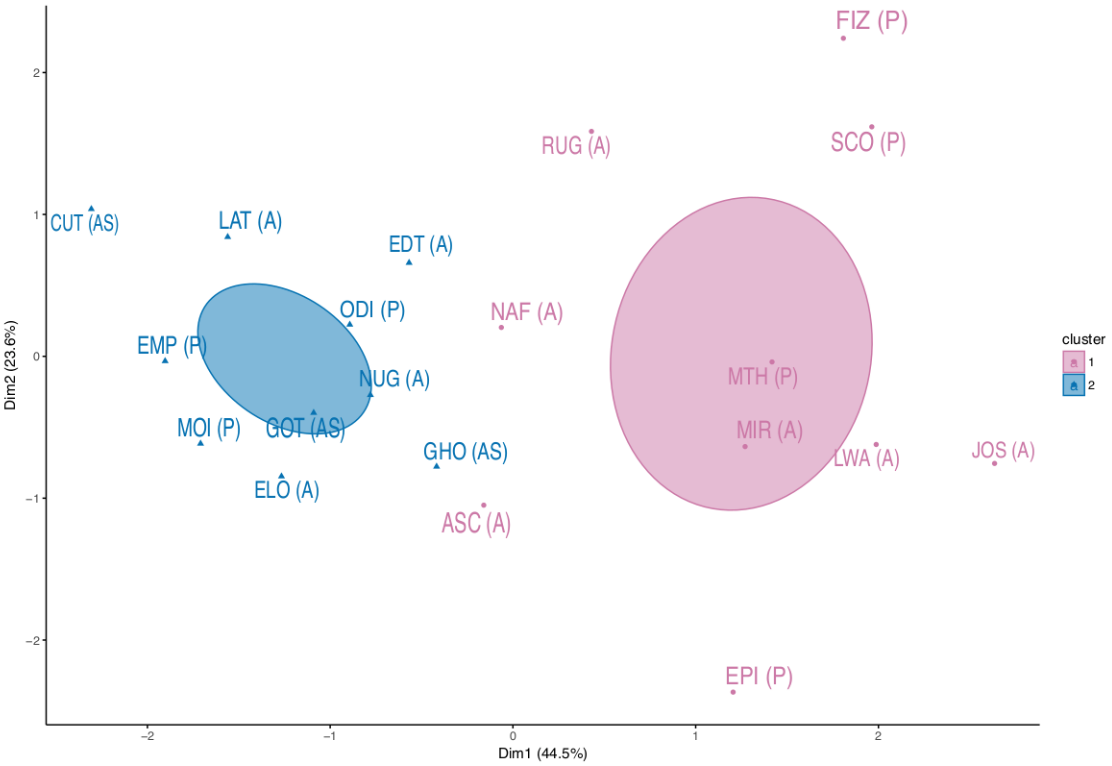
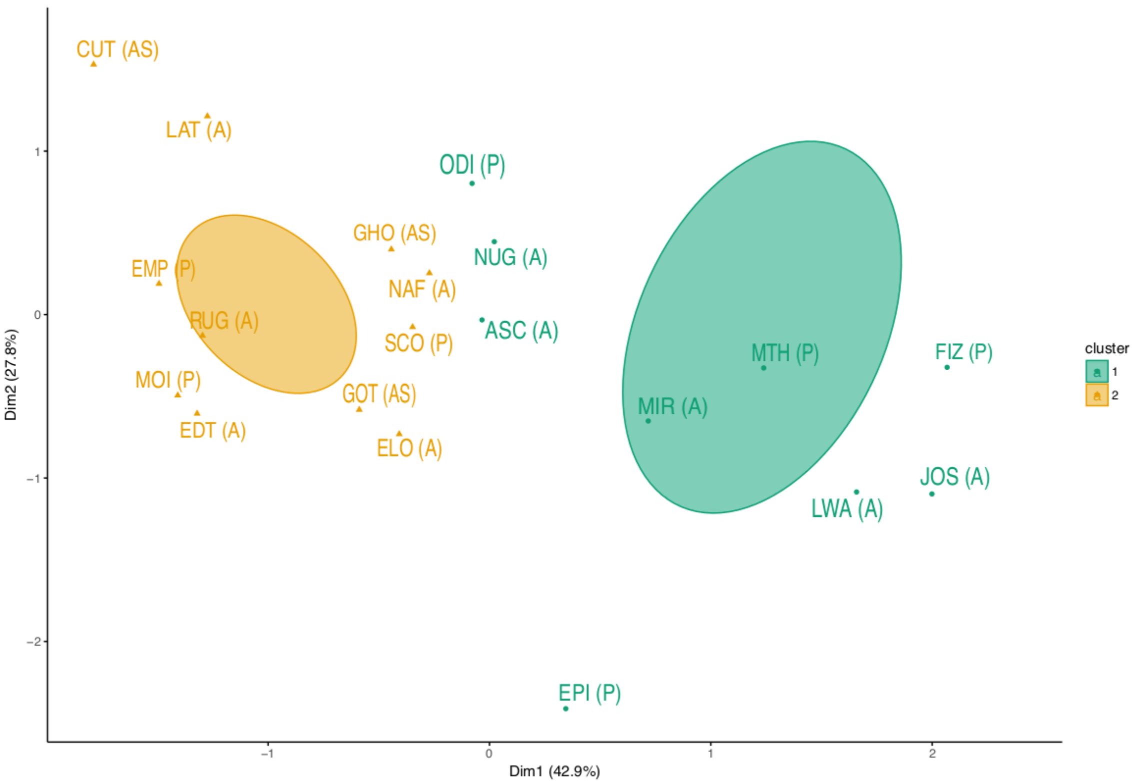


FIGURE 22. CLUSTER ANALYSIS (K-MEANS) ON THE MEASURES OF NWR, BLOCK DESIGN, MATRIX REASONING AND RPM ONLY ON THE FUZZY CLUSTER



Concerning morphosyntactic abilities (Figure 21), the children of the fuzzy cluster were divided into two subgroups: the first cluster incorporated children who displayed impaired performance on SR (*n* = 10), while the second cluster included children who displayed LN performance on SR (*n* = 10). Among these LN children there were two children who performed slightly below the cut-off for impairment (< 78%, Tuller et al. 2018). However, slightly low performance of these two children was due exclusively to some errors (such as lexical substitutions and/or addition of subject clitics) that were not morphosyntactic in nature. As explained in Chapter 5, if we use the *target structure score* of SR, excluding errors as substitutions, both children performed in the norms (GHO = 90 %; EDT = 83%). This did not hold true for children in the LI cluster, whose performance on target structure score remained severely impaired. YAT from the ASD-LN with low NVIQ cluster should be added to the ASD-LI cluster, due to his very low performance on the SR task.

Cluster analysis on NWR resulted in two groups (Figure 22), the first one including all children who displayed impaired performance (*n* = 9) and the second group embodying all children (*n* =11) with LN performance on the task.

In conclusion, the fuzzy cluster was composed by 7 children who displayed LI performance on both SR and NWR, 6 children who displayed LN performance on both SR and NWR and 7 children who were selectively LI either on SR (*n* = 5) or NWR (*n* = 2). The 6 ASD-LN children were not integrated in the ASD-LN with high average/average NVIQ profile detected in previous cluster analyses (PAM and Model-based) because they generally had a lower NVIQ.

6.3.4 Conclusions

Among our ASD group, four profiles of morphosyntactic/NVIQ abilities and four profiles of phonological/NVIQ abilities emerged:

1. ASD-LN with high average/average NVIQ
2. ASD-LN with low /low average NVIQ
3. ASD-LN with average NVIQ
4. ASD-LI with low average/average NVIQ

Profiles 1 and 2 corresponded to the two well-defined clusters found in the previous PAM and Model-based analyses and they were always composed of the same children. The only exception was YAT who will be included in the ASD-LI profile due to his low morphosyntactic abilities. Profiles 3 and 4 were the ones that emerged from our last analysis on the fuzzy cluster. They consisted in a core group of 13 children who displayed the same profile on both morphosyntactic and phonological abilities (both low or both high) and 7 children who had a mixed profile (LN in one domain and LI in the other domain). Table 27 displays the *four* morphosyntactic/NVIQ ability profiles and Table 28 the *four* phonological/NVIQ ability profiles.

TABLE 27. THE MAIN CHARACTERISTICS OF THE FOUR MORPHOSYNTACTIC/NVIQ PROFILES (MEAN, SD)

	Profile 1 ***n* = 8**	**Profile 2** ***n* = 4**	**Profile 3** ***n* = 10**	**Profile 4** ***n* = 11**
SR (%)	98 (16)	88 (4)	90 (9)	23 (17)
RPM (standard score)	110 (6)	78 (8)	92 (8)	81 (8)
Block Design (standard score)	107 (11)	70 (6)	100 (11)	100 (19)
Matrix Reasoning (standard score)	109 (6)	85 (4)	100 (16)	85 (17)
Morphosyntactic / NVIQ ability profiles	ASD-LN with high average /average NVIQ abilities	ASD-LN with low / low average NVIQ abilities	ASD-LN with average NVIQ abilities	ASD-LI with low average/average NVIQ abilities

TABLE 28. THE MAIN CHARACTERISTICS OF THE FOUR PHONOLOGICAL/NVIQ PROFILES (MEAN, SD)

	Profile 1 ***n* = 8**	**Profile 2** ***n* = 5**	**Profile 3** ***n* = 11**	**Profile 4** ***n* = 9**
NWR (percentage)	91 (5)	93 (3)	90 (5)	68 (12)
RPM (standard score)	110 (6)	78 (7)	90 (9)	85 (9)
Block Design (standard score)	107 (11)	69 (9)	100 (12)	100 (15)
Matrix Reasoning (standard score)	109 (6)	85 (13)	95 (19)	95 (13)
Phonological / NVIQ ability profiles	ASD-LN with high average /average NVIQ abilities	ASD-LN with low / low average NVIQ abilities	ASD-LN with average NVIQ abilities	ASD-LI with low average/average NVIQ abilities

Finally, it is worth mentioning that the results of the clustering analyses did not group children according to their diagnostic subcategory. As shown in each cluster analysis, children with Autistic Disorder (A) and PDD-NOS (P) were found in all clusters, and children with Asperger Syndrome (AS) did not display the LI profile.

6.4 Children assessed with the EDEI-R psychometric test

Four children were assessed via the EDEI-R psychometric test and for statistical reasons they were not included in the previous analysis. The EDEI-R battery provides a NV score, but within the battery no test is directly comparable to Block Design and Matrix Reasoning. However, after detecting four profiles of morphosyntactic/NV abilities and four profiles of phonological/NV abilities, we were able to incorporate these four children into these profiles, on the basis of their scores for SR, NWR, RPM and the NV score of EDEI-R. Looking at individual scores (Table 24), ROS could be included in profile 3 (ASD-LN with average NVIQ) and SIM, KEV and EVA in profile 4 (ASD-LI with low average/average NVIQ) for both morphosyntactic and phonological abilities.

TABLE 29. INDIVIDUAL SCORES OF FOUR CHILDREN ASSESSED VIA THE EDEI-R BATTERY

Child code	NWR (%)	SR (%)	RPM (standard score)	NV score (standard score)
SIM (A)	62	40	114	104
KEV (P)	28	33	78	97
EVA (P)	38	33	95	80
ROS (A)	96	90	90	95

Note: Scores below the thresholds are highlighted in grey

6.5 General conclusions and discussion

In this chapter we investigated structural language/nonverbal ability profiles in children with ASD, placing heterogeneity at the centre of our investigation. We used measures that proved to be particularly domain-specific, i.e. structural language (SR and NWR) and NVIQ (RPM, Block Design and Matrix Reasoning), from the analysis in Chapter 5, to the subgroup of 33 children with ASD who were assessed via the WISC-IV, in order to see whether clear profiles would emerge via a statistical approach based on cluster analysis.

After demonstrating that dividing children only on structural language abilities or NVIQ abilities did not provide a clear picture of the profiles existing in our group of children with ASD (Figures 16 and 17), we performed a series of integrative cluster analyses that took into consideration both linguistic and cognitive abilities (Figures 18 and 20). This brought to light two clear profiles, each of which involved normal structural language abilities. One profile (*n* = 8), which corresponds to a group frequently reported on in the literature, combined normal structural language and normal NVIQ (the so-called ASD-LN group in previous studies). The other profile (*n* = 5), which has rarely been reported on in the literature, consisted of normal structural language with low NVIQ. The rarity of this profile in previous studies is at least in part due to the rarity of language studies in ASD including children with low NVIQ. These two profiles concerned approximately 40% of the children in our sample. We note that the discrepant ASD-LN with borderline/low NVIQ profile constituted only 15% of our sample, similarly to what Joseph at al. (2002) and Kjelgaard and Tager-Flusberg (2001) found. It remains to be seen whether this low frequency reflects a true characteristic of the population or whether it is related to the small number of individuals with low NVIQ that participated in our study (*n* = 5), and, more generally, in studies on language in ASD. From the vantage point of the existence of a language module in the human mind/brain, which thus can be selectively spared (see Smith & Tsimpli, 1995), and which nonetheless interfaces with other modules and central systems, our results receive a natural interpretation. The existence of a double dissociation like the one found in the ASD-LI with average NVIQ profile and the ASD-LN with low NVIQ profile indicates that children with ASD can indeed display impaired language abilities in presence of spared nonverbal intelligence, as what is found in SLI, or spared language abilities in the presence of impaired nonverbal intelligence, a profile reminiscent of that found in Williams Syndrome (Mervis & Velleman, 2011) and also in the language Savant *Christopher* (Smith & Tsimpli, 1995).

Moving to the participants who did not fall into these two clear profiles (*n* = 20), most of these children showed discrepant performance in NVIQ abilities and/or in structural language abilities. Regarding NVIQ, some children performed much worse on Matrix Reasoning than on the other two tests, and other children, contra Nader et al. (2016), performed much worse on RPM than on the other two tests. Likewise, regarding structural language abilities, some children had selectively impaired phonology (as found also by Rapin et al. 2009) and others had selectively impaired morphosyntax (a profile frequently claimed to be predominant in ASD; see, for example, Boucher, 2003). Finally, a few children in this fuzzy cluster, despite having uniformly spared or impaired performance within NVIQ and within structural language, displayed some differences of performance on NV measures (at least 10 standard scores of difference between one score and the others).

For these reasons we decided to run another cluster analyses concentrating only on this fuzzy group, separating performance on morphosyntax and phonology, in order to see if the children could be better distinguished. Results showed that two clear profiles emerged once we separated performance on SR and NWR. Within the fuzzy cluster there was a core group of 13 children who displayed the same profile on both morphosyntactic and phonological abilities (both LI, *n* = 5, or LN, *n* = 6) and 7 children who changed ability profile (LI or LN) on the basis of their performance on SR and NWR. Interestingly all three DSM-IV diagnostic subtypes were found in each of the profiles of structural language and NVIQ abilities, except for children with Asperger’s syndrome, who did not show LI profiles. This result would appear to be in accord with the transition to the DSM-5 and the ICD-11 idea of a *spectrum*.

In conclusion, this chapter has shown that normal language performance in children with ASD does not seem to be attributable to normal cognitive abilities. Likewise low NVIQ does not appear to entail impaired language performance. Our analysis found two clear normal language profiles: one including children with normal NVIQ and one with low NVIQ. On the other hand, children with impaired language did not neatly divide between those with normal and impaired NVIQ; in particular, the so-called “SLI” phenotype in ASD did not clearly emerge, since it was obscured by the heterogeneous performance of some children in the fuzzy cluster (n = 14). Only once we separated children with ASD on the basis of their performance on morphosyntax and phonology did a profile similar to the one displayed by SLI emerge, the ASD-LI with low average/average NVIQ. Likewise, no clear ASD-LI with low NVIQ profile, displayed by only one child on both morphosyntax and phonology, emerged for the same reasons. We believe that the findings of this chapter, together, illustrate rather clearly that progress in understanding language profiles in ASD is dependent on wide investigation of the spectrum and use of robust structural language measures.

The next step in our analysis will be to concentrate on structural language abilities and compare the performance of the four profiles of morphosynatctic abilities and the four profiles of phonological abilities with the performance of children with SLI and TD children. This analysis will allow us to investigate the issue of the phenotypical similarities between ASD-LI and SLI and between ASD-LN and TD.

Chapter VII
Computational complexity and error typology on SR and NWR in children with ASD: is there a phenotypical SLI profile?

7.1 Introduction

Results in the literature have shown that ASD-LI and SLI, on the one hand, and ASD-LN and TD, on the other hand, display same overall performance. However some studies reporting on qualitative error analysis of SR and NWR tasks have revealed that these groups may differ on error typology (Chapter 1). The main question we asked here was whether these differences are genuine or not. In this chapter we will try to answer this question through the use of two tasks designed to focus on computational complexity in both morphosyntax and phonology. We will compare children with ASD to children with SLI and TD children with the aim of investigating **whether the phenotypical profile of children with ASD-LI resembles the one shown by children with SLI and whether the phenotypical profile of children with ASD-LN resembles the one displayed by TD children**. We will base our analysis on the four profiles of structural language / NV abilities obtained in Chapter 6. Since these profiles have been controlled for internal homogeneity, the comparison with SLI and TD groups should be "cleaner" than such comparisons typically are.

7.2 Methods

Children will be compared for their performance on SR and NWR in order to see if those having an LI profile resemble children with SLI and if children having an LN profile display similar language abilities as TD children. In order to verify these hypotheses, we will first concentrate on global results, then we will move to qualitative error analysis and finally we will compare developmental trajectories. We will first report results on SR task and then on NWR task.

7.3 The SR task

7.3.1 Participants

The participants were the 37 children with ASD divided into the four profiles of morphosyntactic/NVIQ abilities obtained in Chapter 6. The children with ASD were compared to the group of 26 children with SLI aged 6-12, and to the two control groups of 42 TD children aged 4-5 (TD4-5) and 42 TD children aged 6-12 (TD6-12). Having both younger and age-matched children allowed to better compare linguistic outcomes of children with ASD in a developmental perspective.

7.3.2 Data analysis

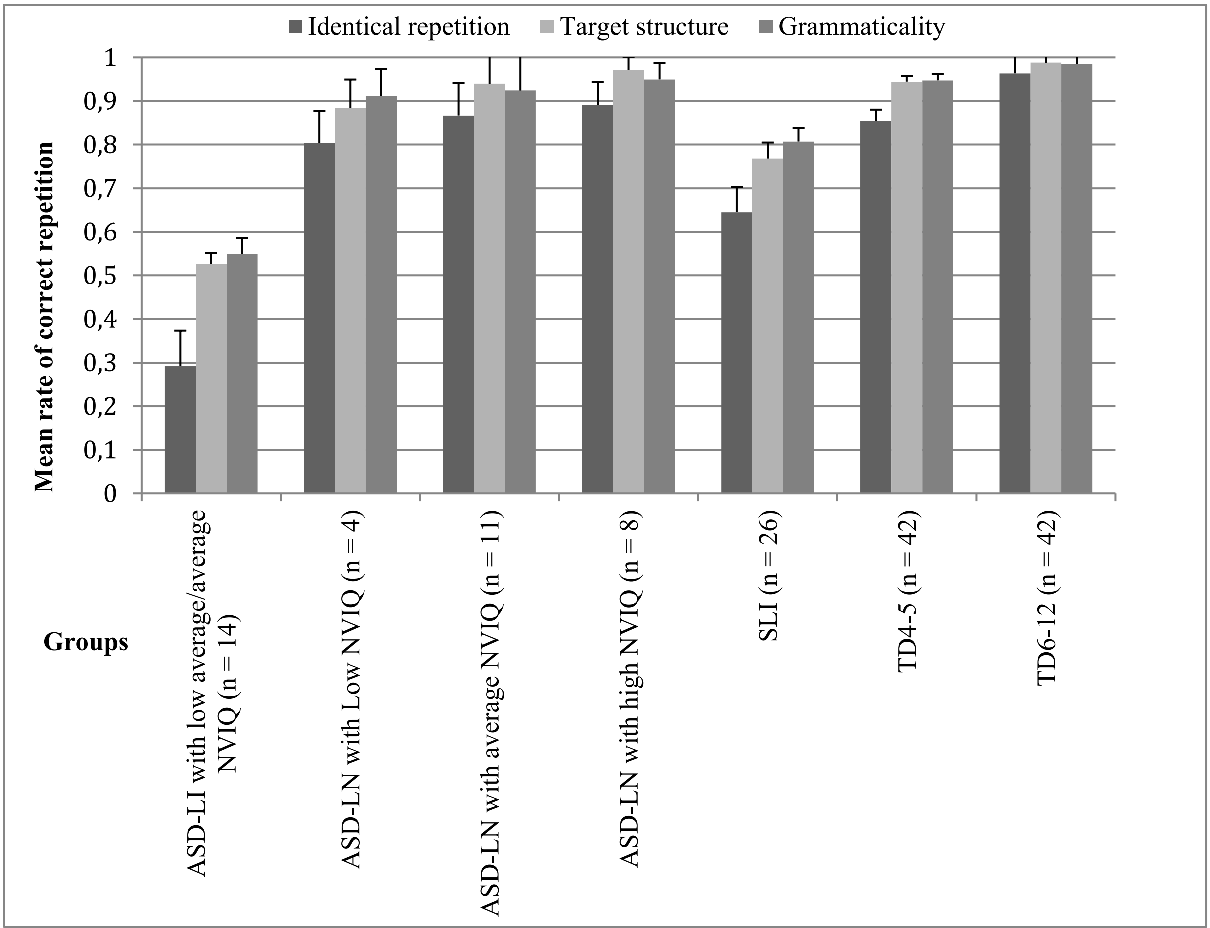
Due to the non-normal distribution of the data (confirmed by the Shapiro-Wilk test) and the small number of children in each subgroup (profile) of children with ASD, our analyses were conducted with nonparametric tests. Post-hoc t-tests included Bonferroni corrections in order to avoid type 1 errors due to multiple comparisons. Due to the very low number of children in some ASD profiles, we used the (Crawford et al., 2010) Singlims program, which includes a t-test for comparison of a single case to a control population. This allowed us to establish, for each individual with ASD, how similar his/her performance was compared to the control groups.

7.3.3 Results on SR task

7.3.3.1 Global results

We first report global results of the six groups on the SR task on three measures: identical repetition, grammaticality and target structure (see Figure 23). Recall that *Identical Repetition* stood for verbatim repetitions of the stimulus sentence, *Grammaticality* for production of a grammatical sentence independently from the targeted structure and *Target structure* for the preservation of the structure targeted in the stimulus sentence even in presence of other errors (substitutions or omissions).

FIGURE 23. MEAN PRODUCTION RATES FOR IDENTICAL REPETITION, GRAMMATICALITY AND TARGET STRUCTURE IN EACH GROUP



Significant intra-group differences were found in almost every group between mean rates of Identical repetition and Target structure and between Identical repetition and Grammaticality (see Table 30). No significant difference was found between Target structure and Grammaticality. The significance threshold after post-hoc correction – Bonferroni was $p < .025$.

Since the Grammaticality score did not differ from the Target structure score, we decided to exclude it from further analyses in order to avoid redundant effects. Moreover, looking at how well children were able to repeat targeted structures of different levels of complexity provides a way of examining the effect of computational complexity more directly than looking merely at whether the sentence the child produced was grammatical or not.

Table 30. Differences between mean rates for Identical repetition, Target structure and Grammaticality in the seven groups

Groups		Identical repetition / Target structure	Identical repetition / Grammaticality	Grammaticality / Target structure
ASD-LI with low average / average NVIQ (n = 14)	Z	-3.296	-3.297	-.817
	p	**.001**	**.001**	.414
ASD-LN with low NVIQ (n = 4)	Z	-1.826	-1.841	-1.633
	p	.068	.066	.102
ASD-LN with average NVIQ (n =11)	Z	-2.812	-2.812	-1.725
	p	**.005**	**.005**	.084
ASD-LN with high NVIQ (n = 8)	Z	-1.826	-1.841	-1.604
	p	.068	.066	.109
SLI (n =26)	Z	-3.534	-3.691	-1.848
	p	**<.001**	**<.001**	.065
TD4-5 (n = 42)	Z	-5.349	-5.551	-.313
	p	**<.001**	**<.001**	.754
TD6-12 (n = 42)	Z	-4.010	-3.684	-1.508
	p	**<.001**	**<.001**	.132

Note: Significant differences are in boldface

Moving to inter-group comparisons, significant between-group differences were found on Identical repetition (X^2 (6, N = 147) = 666.818, p < .001) and Target structure (X^2 (6, N = 147) = 2185.861, p < .001). Mann-Whitney intergroup comparisons (Post-hoc correction p < .008) showed that the ASD-LI group performed worse than any other group of children, including the SLI group. The three ASD-LN groups did not differ from each other. While the ASD-LN with high NVIQ group performed like the TD4-5 and TD6-12 groups, the ASD-LN with low NVIQ and ASD-LN with average NVIQ groups performed respectively more like the SLI and TD4-5 groups, and they performed significantly lower than the TD6-12 group. Finally, the SLI group performed significantly lower than the TD4-5 and TD6-12 groups and the TD4-5 group performed significantly lower than TD6-12 group. Intergroup comparisons are presented in Table 31, and general group performance is shown in Figure 24 for Identical repetition and in Figure 25 for Target structure.

TABLE 31. MANN-WHITNEY INTERGROUP COMPARISONS BETWEEN THE SEVEN GROUPS ON THE MEASURES OF IDENTICAL REPETITION AND TARGET STRUCTURES

Inter-group comparisons	Identical repetition			Target structure		
	U	*p*	*r*	*U*	*p*	*r*
ASD-LN with High NVIQ (*n*= 8) / ASD-LN with low NVIQ (*n*= 4)	10	.299	- .28	10.5	.315	- .15
ASD-LN with High NVIQ (*n*= 8) / ASD-LI with low average - average NVIQ (*n*= 14)	0	< **.001**	- .53	0	< **.001**	- .52
ASD-LN with High NVIQ (*n*= 8) / ASD-LN with average NVIQ (*n*= 11)	30.5	.259	- .26	32.5	.302	- .24
ASD-LN with High NVIQ (*n*= 8) / SLI (*n*= 26)	39	.008	- .45	39.5	.008	- .45
ASD-LN with High NVIQ (*n*= 8) / TD4-5 (*n*= 42)	102	.077	- .25	136	.373	- .12
ASD-LN with High NVIQ (*n*= 8) / TD6-12 (*n*= 42)	157	.758	- .04	134.5	.236	- .16
ASD-LN with low NVIQ (*n*= 4) / ASD-LI with low average - average NVIQ (*n*= 14)	0	**.003**	- .76	3	**.003**	- .69
ASD-LN with low NVIQ (*n*= 4) / ASD-LN with average NVIQ (*n*=11)	20.5	.844	- .01	20.5	.838	- .02
ASD-LN with low NVIQ (*n*= 4) / SLI (*n*= 26)	27.5	.134	- .20	29.5	.168	- .16
ASD-LN with low NVIQ (*n*= 4) / TD4-5 (*n*= 42)	69	.526	- .18	65.5	.454	- .13
ASD-LN with low NVIQ (*n*= 4) / TD6-12 (*n*= 42)	12	**.003**	- .47	33.5	.011	- .44
ASD-LI with low average - average NVIQ (*n*= 14) / ASD-LN with average NVIQ (*n*= 11)	0	< **.001**	- .86	0	< **.001**	- .86
ASD-LI with low average - average NVIQ (*n*= 14) / SLI (*n*= 26)	38.5	< **.001**	- .62	65	**.002**	- .50
ASD-LI with low average - average NVIQ (*n*= 14) / TD4-5 (*n*= 42)	1	< **.001**	- .73	6	< **.001**	- .72
ASD-LI with low average - average NVIQ (*n*= 14) / TD6-12 (*n*= 42)	0	< **.001**	- .74	0	< **.001**	- .82
ASD-LN with average NVIQ (*n*= 11) / SLI (*n*= 26)	70	.015	- .40	72.5	.018	- .39
ASD-LN with average NVIQ (*n*= 11) / TD 4-5 (*n*= 42)	221.5	.833	- .02	215.5	.723	- .04
ASD-LN with average NVIQ (*n*= 11) / TD 6-12 (*n*= 42)	102.5	**.004**	- .40	133.5	**.007**	- .37
SLI (*n*= 26) / TD4-5 (*n*= 42)	286	< **.001**	- .40	267	< **.001**	- .43
SLI (*n*= 26) / TD6-12 (*n*= 42)	99	< **.001**	- .70	147.5	< **.001**	- .67
TD4-5 (*n*= 42) / TD6-12(*n*= 42)	251.5	< **.001**	- .70	515.5	< **.001**	- .40

FIGURE 24. MEAN RATE OF CORRECT IDENTICAL REPETITION IN EACH GROUP

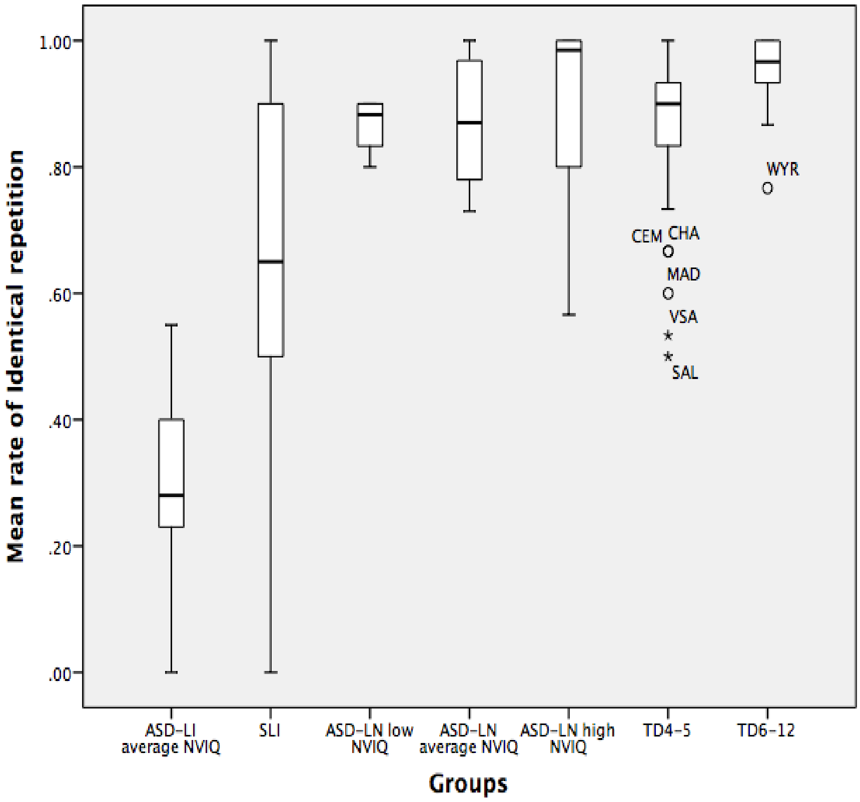
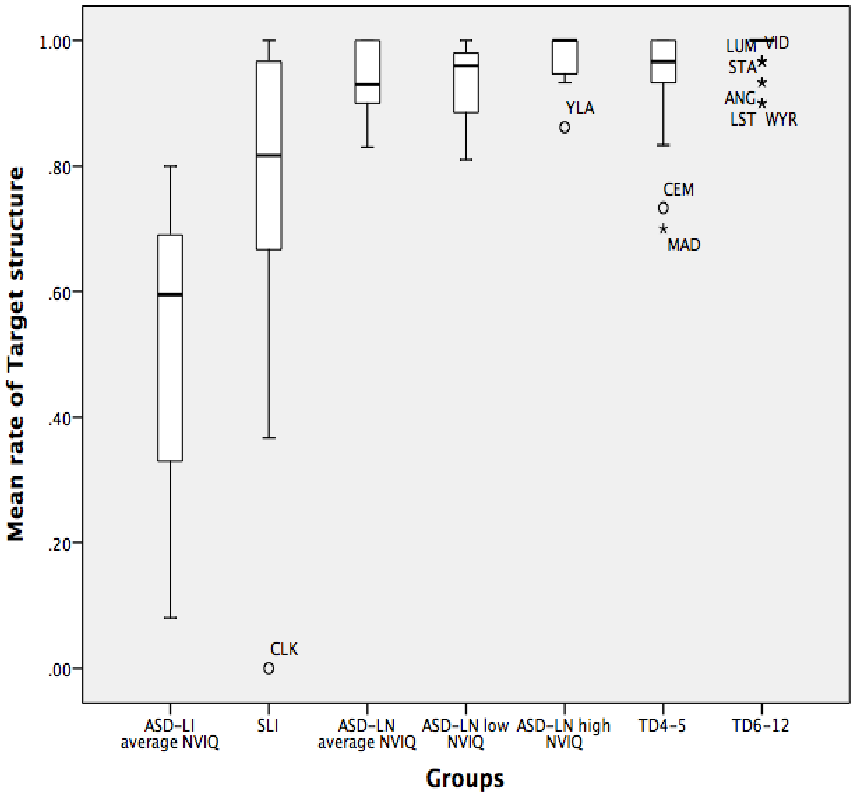


FIGURE 25. MEAN RATE OF CORRECT TARGET STRUCTURE IN EACH GROUP



Note: the cut-off for impaired performance was established at 78%

Global results (Table 31 and Figures 24) on Identical repetition suggested that the ASD-LI group performed much worse than all the other groups, including the SLI group (except for one child “CLK”). While no child with ASD-LI performed over 55% correct identical repetition, the SLI group was much more heterogeneous, ranging from zero correct repetition to ceiling performance. These results may be interpreted as unusual for children with SLI, who are generally expected to show significant effects of computational complexity. However, in our SLI group older children (9- to 12-years old) performed significantly better than younger children with SLI (6- to 8- years old), rising mean group performance. We will come back to this topic in section 7.3.3.5. Moreover, although Target structure scores were generally higher than Identical repetition scores, a significant difference between ASD-LI and SLI performance was still found, indicating that somehow language impairment in ASD-LI was much severe than SLI.

Concerning the ASD-LN and TD groups, except for the children in the ASD-LN with high NVIQ group, who performed like age-matched TD peers (TD6-12), the other two ASD-LN subgroups did not differ from the TD4-5 and SLI groups (Table 31). However, Figures 24 and 25 show that group performance in children with ASD-LN more strongly resembled that of the TD4-5 than that of the children with SLI. The fact that some groups, like the ASD-LN with low NVIQ group, included only 4 children might have led to high Type 1 errors, such as low group score linked to the particular performance of one or two individuals in rather small groups. We decided then to look at individual performance to see whether individual scores replicated tendencies displayed by global results.

7.3.3.2 Individual performance

We used the Crawford et al., (2010) Singlims program, which includes a *t*-test for comparison of a single case to a control population. We compared the performance of each child in the ASD-LI group with the performance of the SLI group (Table 32), and the performance of each child in the three ASD-LN profiles with the performance of the two control groups (Table 33).

TABLE 32. INDIVIDUAL PERFORMANCE OF CHILDREN IN THE ASD-LI GROUP COMPARED WITH THE PERFORMANCE OF THE SLI GROUP, USING THE CRAWFORD T-TEST

	Identical repetition		**Target structure**	
Child Code	*t*	*p*	*t*	*p*
FIZ	-0.528	.300	-0.555	.292
SIM	-0.906	.186	-1.280	.106
SCO	-2.146	.011	-2.901	.003
RUG	-1.812	.041	0	.500
ASC	-1.547	.067	-2.133	.021
KEV	-1.170	.126	-0.853	.401
MTH	-1.170	.126	-0.555	.292
LWA	-1.547	.067	-1.408	.085
EVA	-1.170	.126	0.171	.043
NAF	-0.528	.300	0	.500
EPI	-1.547	.067	-1.835	.039
JOS	-2.189	.019	-1.963	.030
MIR	-1.547	.067	-0.384	.352
YAT	-0.340	.368	-0.264	.396

Note: grey cells indicate significant differences

TABLE 33. INDIVIDUAL PERFORMANCE OF CHILDREN IN THE ASD-LN PROFILES COMPARED WITH THE PERFORMANCE OF THE CONTROL GROUPS TD4-5 AND TD6-12, USING THE CRAWFORD T-TEST

		TD4-5				TD6-12			
		Identical repetition		Target structure		Identical repetition		Target structure	
	Child Code	*t*	*p*	*t*	*p*	*t*	*p*	*t*	*p*
ASD-LN with High NVIQ	MON	-1.078	.143	-0.141	.444	-4.546	<.001	-2.471	.008
	LEC	1.348	.092	0.847	.200	0.791	.216	0.988	.164
	YLA	-2.606	.006	-1.129	.132	-7.906	<.001	-5.930	<.001
	TUC	1.348	.092	0.847	.200	0.791	.216	0.988	.164
	NOS	0.180	.429	0.282	.389	-1.779	.041	-0.988	.164
	MUG	1.078	.143	0.847	.200	0.198	.422	0.988	.164
	VOR	1.348	.092	0.847	.200	0.791	.216	0.988	.164
	AVI	1.348	.092	0.847	.200	0.791	.216	0.988	.164
ASD-LN with low NVIQ	ADO	-0.449	.327	-1.835	.036	-3.163	.001	-8.041	<.001
	ATE	0.449	.327	0.282	.389	-1.186	.121	-0.988	.164
	LPG	0.449	.327	-1.835	.039	-1.186	.121	0.988	.164
	CIP	0.180	.429	0.282	.389	-1.977	.027	-0.988	.164
ASD-LN with average NVIQ	ODI	-0.180	.429	-0.565	.287	-2.570	.006	-3.953	<.001
	CUT	1.348	.092	0.847	.200	0.791	.216	0.988	.164
	NUG	1.078	.143	0.847	.200	0.198	.422	0.988	.164
	ROS	-1.078	.143	-0.565	.287	-4.546	<.001	-3.953	<.001

EDT	-1.078	.143	-1.553	.064	-4.546	<.001	-7.412	<.001
EMP	0.988	.164	0.847	.200	0	.500	0.988	.164
GOT	1.078	.143	0.847	.200	0.198	.422	0.988	.164
ELO	0.180	.429	-0.565	.287	-1.779	.041	-3.953	<.001
GHO	-0.809	.211	-0.565	.287	-3.953	<.001	-3.953	<.001
LAT	-0.449	.327	-0.141	.444	-3.163	.001	-2.471	.008
MOI	0.719	.238	0.847	.200	-0.593	.278	0.988	.164

Note: grey cells indicate significant differences

The results in Table 32 indicate that 8/14 children with the ASD-LI profile performed in line with the SLI group, while 6/14 children performed significantly worse on one or both scores. Table 33 indicates that children in the ASD-LN subgroups tended to perform more often like the TD4-5 group than like the TD6-12 group. This was evident in the ASD-LN average NVIQ profile, where 6/11 children showed significant differences with the TD6-12 group, but no child differed from the TD4-5 group. In the ASD-LN high NVIQ profile, children performed like both the TD4-5 and TD6-12 groups, with the exception of two children who performed significantly below the TD6-12 group on both Identical repetition and Target structure measures. As already explained in Chapter 6, the drop of performance of these two children ("YLA" and "MON") was due to a lack of collaboration and/or attention during the testing phase. Concerning the ASD-LN with low NVIQ group, individual performance was generally in line with both the TD4-5 and TD6-12 groups except for one child ("ADO"), who performed significantly below both groups (more markedly with respect to the TD6-12 group).

In general we can see that the performance found in group results was only partially confirmed by individual results. Individual performance suggested that not every child with ASD-LI performed significantly lower than the SLI group, and not every child with ASD-LN with high NVIQ performed like his/her age peers in the TD6-12 group. Nonetheless, although statistical analysis highlighted significant differences between the ASD-LI and the SLI group only for 6/14 children with ASD-LI, none of the 14 children with ASD-LI performed like or better than the SLI group (on both scores). This means that children with ASD-LI systematically displayed lower scores than the SLI group. For the ASD-LN groups all children performed in line with the TD4-5 group, while only 12/23 showed performance in line with the TD6-12 group.

Although this first analyses (both as group performance and individual performance) yielded some interesting results, we hypothesised that looking at global scores of Identical repetition and Target structure was not sufficient to properly measure the gradient of similarities and differences between the ASD-LI and SLI groups and between the ASD-LN and TD groups. Since SR included sentences at different level of complexity, it could be assumed that if children were sensitive to computational complexity, they would not perform homogeneously on less complex and more complex substructures. In order to investigate this assumption, we opted for a much more detailed analysis that took into consideration the effect of computational complexity in the SR task. It could be the case, in fact, that low general scores of Identical repetition and Target structure in the ASD-LI group were the results of a dramatic drop of performance on one (or more) substructures, which differed in computational complexity, that may have obscured fundamental similarities between children with ASD-LI and children with SLI.

7.3.3.3 Performance on substructures: the role of computational complexity

We searched for possible effects of computational complexity in children's performance. Generally, we hypothesised that the language impaired groups, ASD-LI and SLI, should perform significantly worse than the ASD-LN and TD groups. For the language impaired groups, we hypothesized that if children with ASD-LI and children with SLI were both sensitive to linguistic computational complexity, monoclausal sentences (monoclausal present tense and past tense) should be repeated better than sentences with embedded clauses (argument clauses and relative clauses) and sentences clauses involving movement (object wh-clauses and relative clauses). Since the number of syntactic dependencies (agreement) and/or movements was greater in object wh-questions, relative clauses and argument clauses than in monoclausal sentences (SVO present and past tense), we expected that language impaired children would performed significantly lower on those conditions that demand a much greater calculation of computational complexity than on the simple SVO sentences. Moreover within each condition, children with language impairment (ASD-LI and SLI) were expected to display effects of computational complexity in their performance on all five substructures, repeating less complex substructures better than the corresponding more complex substructures. Finally, we hypothesised that if the morphosyntactic skills of children with ASD-LI were similar to those of children with SLI, the two groups should perform alike both on global results and on errors types.

Concerning the ASD-LN and TD groups, we hypothesized that both the ASD-LN groups and the TD groups should not display computational complexity effects, except for selective drop of performance on the most complex conditions involving embedded clauses (object relatives and finite complement clauses) for younger TD4-5 children. This hypothesis was based on previous findings in the literature on TD children. Since computational complexity has effects in young TD children, we expected that the TD4-5 group would display higher rates of errors on *quel* wh-object condition and object relative condition, because of RM effects (the presence of a strong intervener has been shown to have an effect at this age) and on finite argument clauses because of increasing computational costs due to the presence of complementizer-tense dependencies and overt subject-verb agreement. On the other hand, TD6-12 children should display significantly less important drop in performance on object relative clauses and on finite argument clauses. Finally, we hypothesised that if morphosyntactic skills of children with ASD-LN were indeed "normal" (not deviant or delayed), these children should perform like their age-matched TD peers.

7.3.3.3.1 Intergroup comparison

The same differences in global performance found between the ASD-LI group and the three subgroups of ASD-LN was confirmed by the analysis of the different structures and substructures. A fine-grained look at performance for Identical repetition and Target structure on the various substructures suggested that the ASD-LI group was more affected than the three ASD-LN groups (Figure 26 and Figure 27).

FIGURE 26. GROUP PERFORMANCE ON IDENTICAL REPETITION OF SUBSTRUCTURES IN THE SR TASK

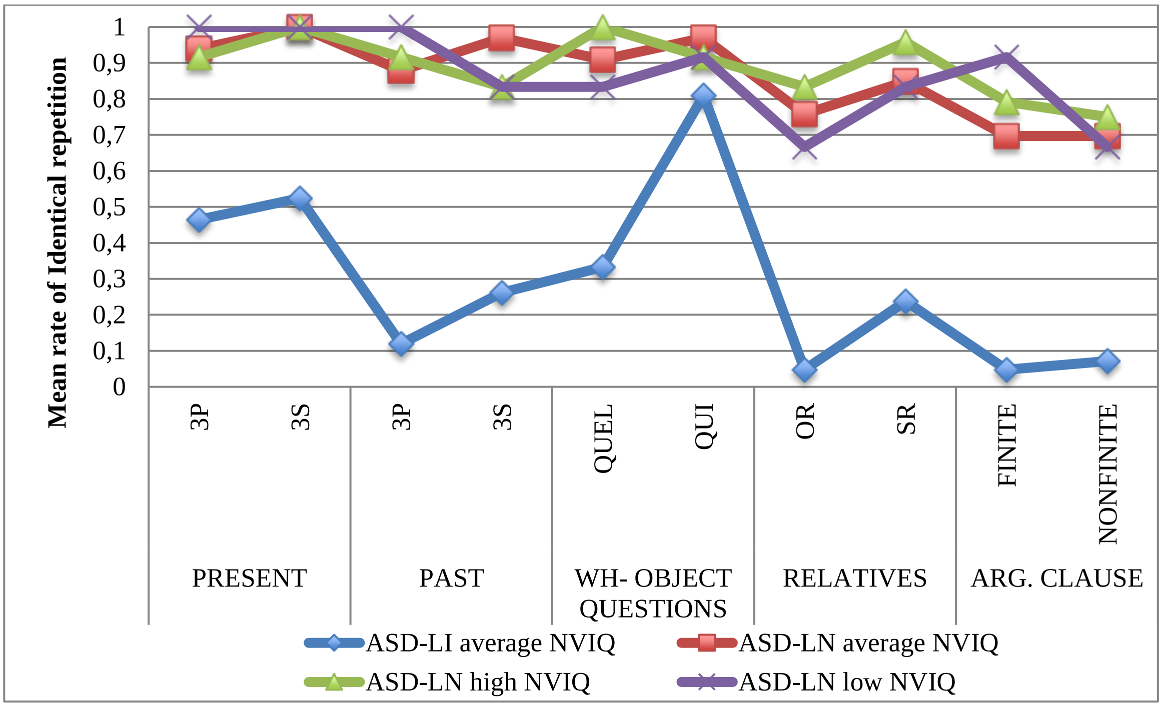
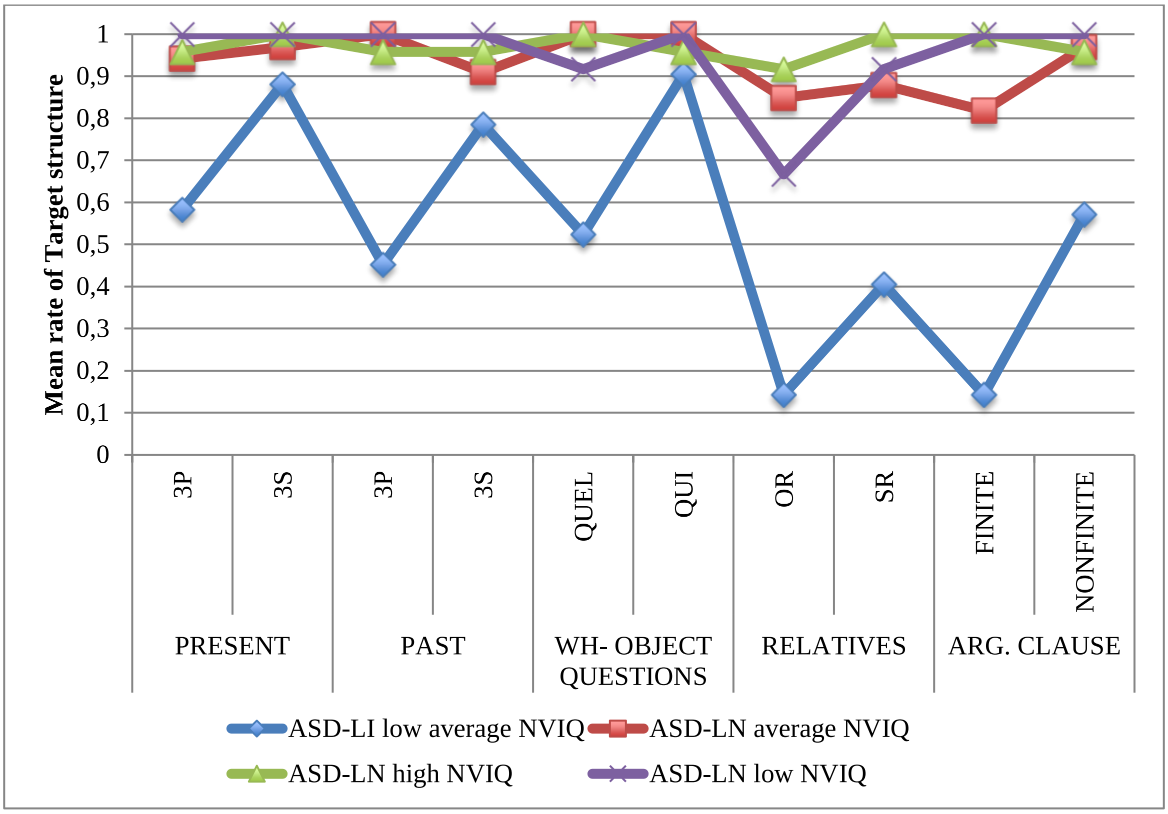


FIGURE 27. GROUP PERFORMANCE ON TARGET STRUCTURE OF SUBSTRUCTURES IN THE SR TASK



Significant differences ($p < .002$ after post-hoc correction) were found between the ASD-LI and ASD-LN with high NVIQ groups on all five conditions for both less complex and more complex substructures, except for Identical repetition of *qui* 'who' object wh-questions ($U(21) = 48$, $p = .506$). Significant differences on all substructures were found on Target structure scores as well, with the exception of present tense 3S ($U(21) = 40$, $p = .104$), past tense 3S ($U(21) = 45.5$, $p = .329$) and *qui* 'who' object wh-questions ($U(21) = 54.5$, $p = .864$).

Comparison between the ASD-LI and ASD-LN with low NVIQ groups exhibited significant differences on each substructure except for *qui* 'who' object wh-questions ($U(17) = 24$, $p = .610$) on Identical repetition. No significant difference emerged between the two groups when compared for Target structure scores, except for finite argument clauses ($U(17) = 2.5$, $p = .001$, $r = -.77$). However a tendency toward significance was found for present tense 3P ($U(17) = 8$, $p = .025$), past tense 3P ($U(17) = 6$, $p = .014$), object relatives ($U(17) = 10$, $p = .029$), subject relatives ($U(17) = 7$, $p = .021$) and nonfinite argument clauses ($U(17) = 8$, $p = .024$).

Finally comparison between the ASD-LI and ASD-LN with average NVIQ groups showed a significant difference on every substructures except for *qui* object wh-questions

($U(24) = 55.5$, $p = .115$) on Identical repetition. Similar results were found on Target structure scores, with the exception of present tense 3S ($U(24) = 55$, $p = .059$), past tense 3S ($U(24) =55$, $p = .059$) and *qui* 'who' object wh-questions ($U(24) = 66$, $p = .201$).

Performance in the ASD-LI group was significantly worse on all substructures than both the TD4-5 and TD6-12 groups, except for *qui* object wh-questions on Target structure score (ASD-LI/TD4-5: $U(61) = 265$, $p = .219$; ASD-LI/TD6-12: $U(61) = 240$, $p = .098$).

Comparison between the three ASD-LN subgroups did not show any significant differences for any substructures of the five conditions, for Identical repetition and Target structure scores.

The three ASD-LN groups differed from the SLI group on object and subject relatives and on finite and nonfinite argument clauses on both Identical repetition and Target structure scores. No other difference was found between the ASD-LN and the SLI groups.

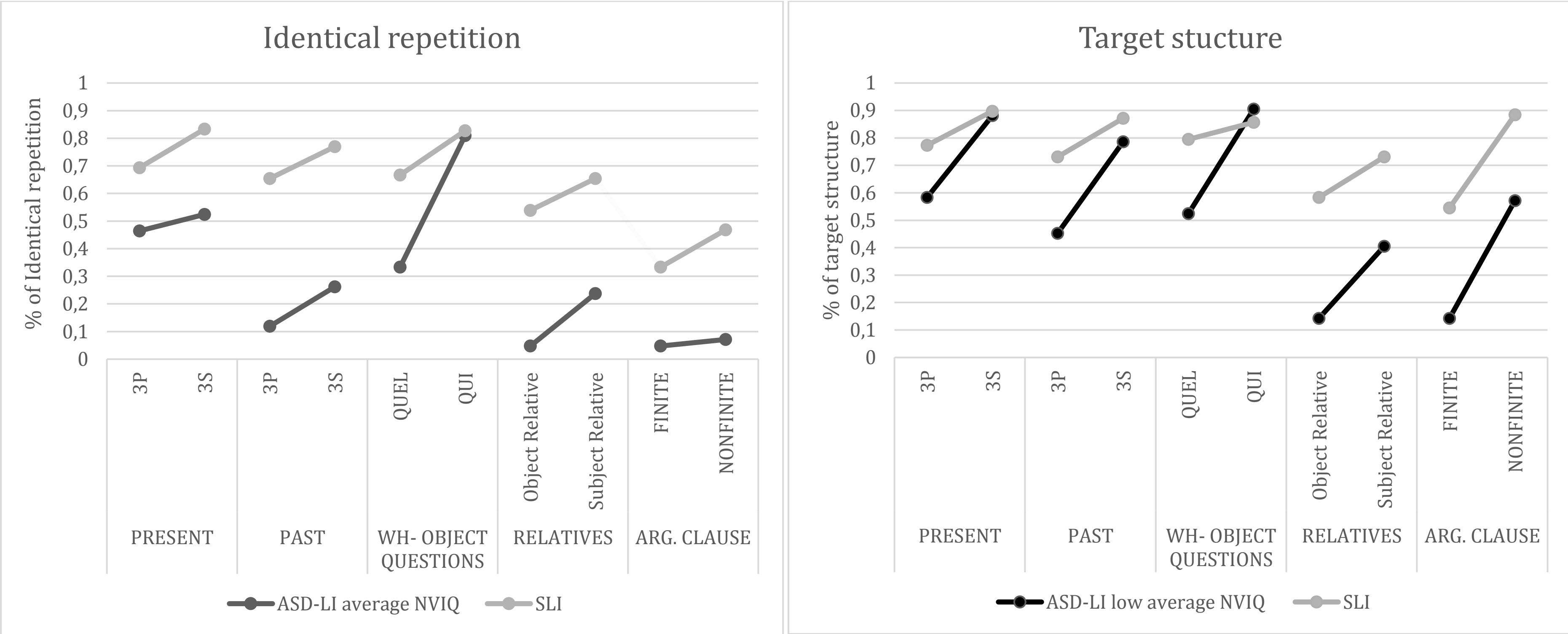
Moving to comparisons across the control groups, the SLI group differed from the TD4-5 on every substructure except for present tense 3S ($U(67) = 441$, $p = .039$) and past tense 3P ($U(67) = 427$, $p = .084$) on Identical repetition score and for past tense 3S ($U(67) = 484$, $p = .112$) and nonfinite argument clauses ($U(67) =431$, $p = .050$) for Target structure score. The SLI group also differed from the TD6-12 group on every substructure, on both Identical repetition and Target structure scores.

Finally, the TD4-5 group performed significantly below the older TD6-12 group on past tense 3P ($U(83) = 598$, $p = .001$, $r = -.37$), object relatives ($U(83) = 570$, $p = .001$, $r = -.35$), finite argument clauses ($U(83) = 467.5$, $p < .001$, $r = -.41$) and nonfinte argument clauses ($U(83) = 418.5$, $p < .001$, $r = -.47$). No significant difference was found when the two TD groups were compared on Target structure scores.

To sum up, the ASD-LI group performed significantly worse than all three ASD-LN groups and than the two TD groups on all substructures (both less and more complex), except for *qui* object wh-questions, on both Identical repetition and Target structure scores. On the Target structure measure, the children with ASD-LI performed like their ASD-LN peers on present tense 3S and past tense 3S, and on *quel* object wh-questions (but only when compared to the ASD-LN with low NVIQ group). The ASD-LN groups did not differ from each other and they performed like the children with SLI on relative and argument clauses. From these results we can draw three conclusions: (1) the *qui* wh-object condition seemed not to cause any problem to any group (including the ASD-LI group); (2) the children with ASD-LI seemed to have the ability to repeat at least less complex monoclausal sentences; (3) the children with ASD-LN seemed to have some trouble repeating sentences with an embedded clause.

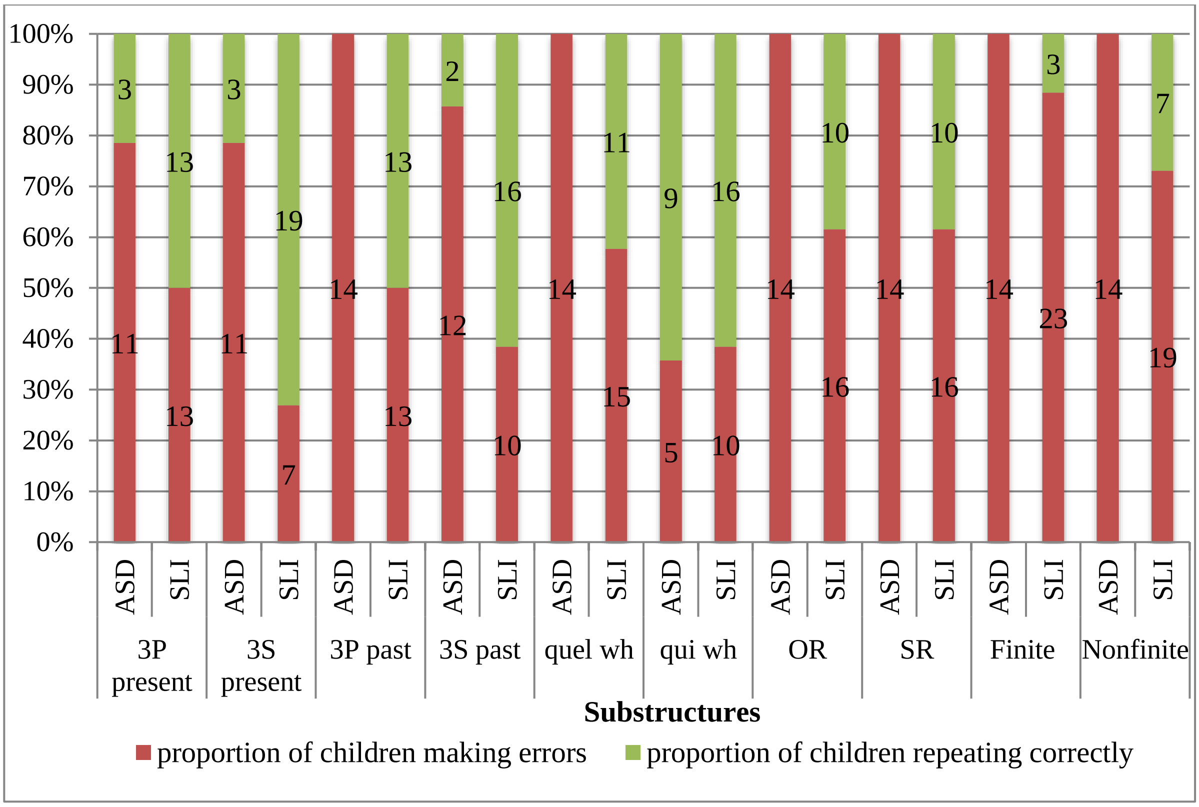
The next step in our analysis was to compare the children with ASD-LI with the children with SLI, and the performance of the children with ASD-LN with the TD children, in order to see if the global results were confirmed by performance on the different substructures. We compared performance for both Identical repetition and Target structure scores. Figure 28 displays the performance of the ASD-LI and SLI groups on the different substructures.

FIGURE 28. ASD-LI AND SLI PERFORMANCE ON SUBSTRUCTURES IN THE SR TASK



The ASD-LI subgroup performed significantly lower on each substructure than the SLI group on Identical repetition scores, except for present tense 3P ($U(39) = 115$, $p = .067$) and *qui* object wh-questions ($U(39) = 181$, $p = .987$). Since the SLI group was almost twice as big as the ASD-LI group, we checked for individual performance in order to verify that these differences were not linked to particular individuals. Figure 29 shows the proportion (and number) of children with 100% correct repetition (in green bars), versus those having less than 100% correct repetition (in red bars), in each group on Identical repetition scores.

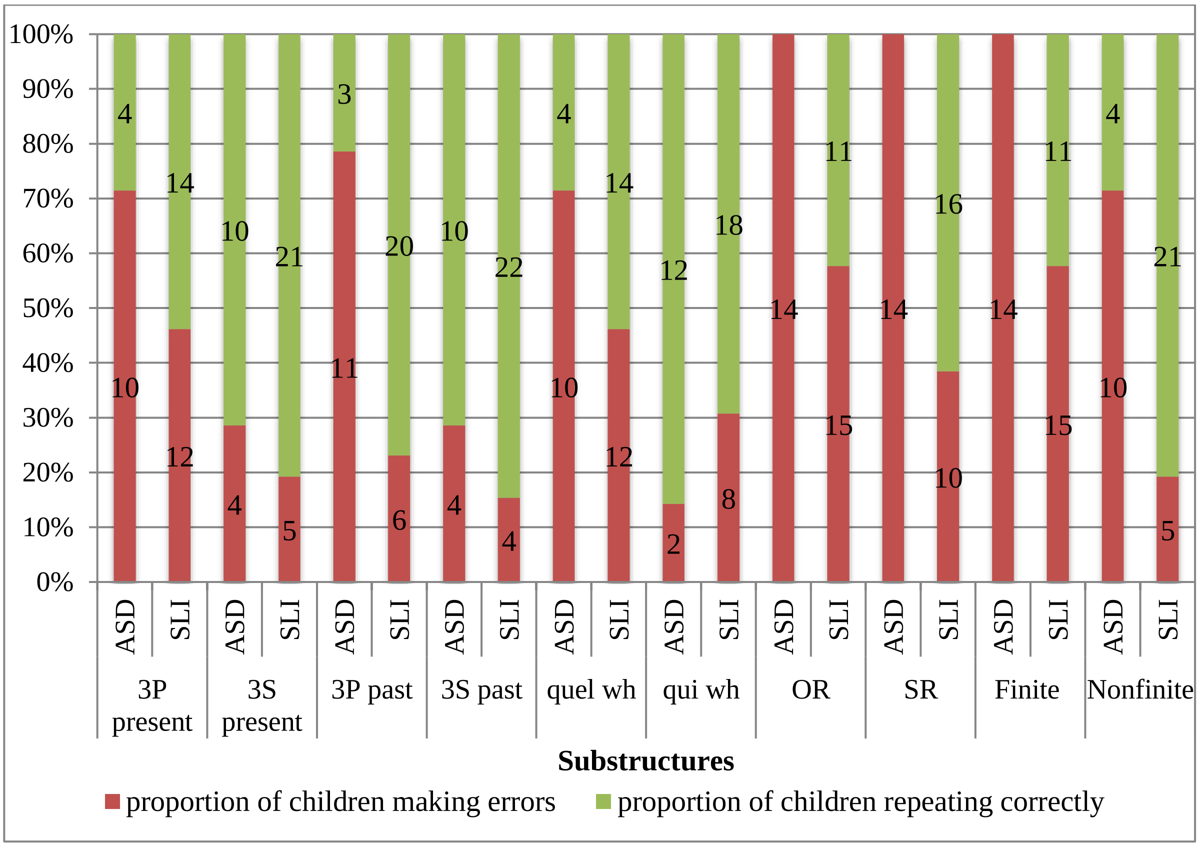
FIGURE 29. PROPORTION (AND NUMBER) OF CHILDREN IN THE ASD-LI AND SLI GROUPS MAKING ERRORS AND CHILDREN REPEATING CORRECTLY ON SUBSTRUCTURES (IDENTICAL REPETITION SCORE)



As can be seen in Figure 29, on six conditions (past tense 3P, *quel* object wh-questions, object and subject relatives, and finite and nonfinte complement clause) the totality of children with ASD-LI (14/14) made errors. Among the SLI group roughly 60% of the children made errors on the same conditions, with the exception of argument clauses (both finite and nonfinite), where more than 75% of the children with SLI performed incorrectly. On the 3S present and 3S past conditions the rate of children with inaccurate repetition performance was twice as big for the ASD-LI group than for the SLI group. The only condition for which the same distribution of children was observed in the two groups was with *qui* wh-object (with roughly 35% of children with less than 100% performance in each group). These results indicate that on average the mean rate of children making errors in our ASD-LI group with respect to Identical repetition was twice as big as the mean rate of children with SLI making errors for all conditions, except for *qui* wh-object.

When compared on Target structure scores, the two groups differed only on object relatives ($U(39) = 78$, $p = .002$) and nonfinite argument clauses ($U(39) = 87.5$, $p = .002$), while a tendency toward significance was found for subject relatives ($U(39) = 96$, $p = .010$) and finite argument clauses ($U(39) = 90$, $p = .006$). In general, then, if we concentrate on Target structure scores, the children with ASD-LI tended to perform lower than the children with SLI on embedded clauses with movement (relative clauses) and without movement (argument clauses). Figure 30 shows the analysis of individual performance for Target structure scores. Again, the number of children involved is detailed for each condition.

FIGURE 30. PROPORTION (AND NUMBER) OF CHILDREN IN THE ASD-LI AND SLI GROUPS MAKING ERRORS AND CHILDREN REPEATING CORRECTLY ON SUBSTRUCTURES (TARGET STRUCTURE SCORE)

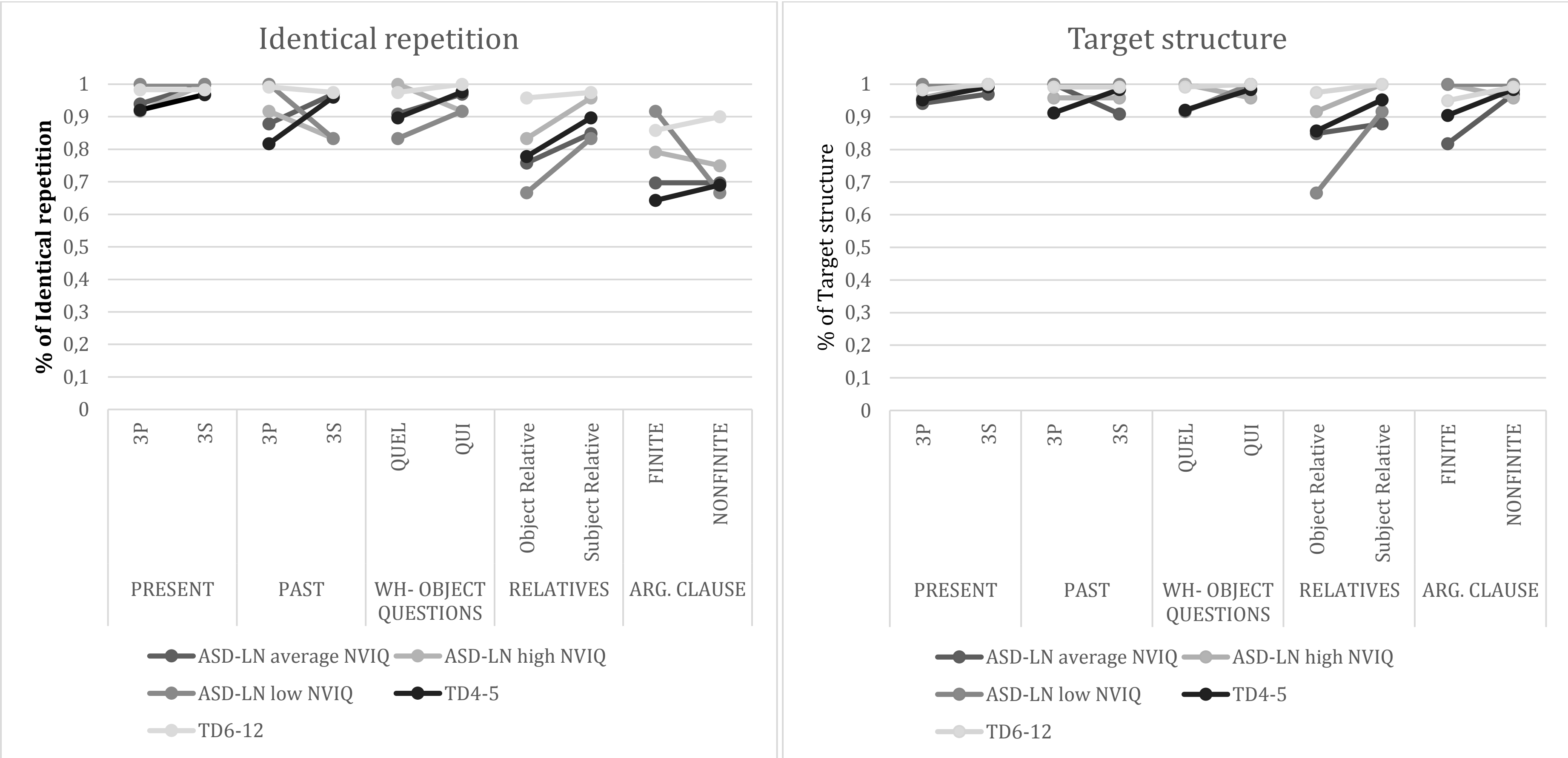


Concerning the number of children making errors in each group for Target structure scores, we notice first of all that while generally for the monoclausal less complex conditions (3S past and present, and *qui* wh-object questions) the children with ASD-LI performed in line with children with SLI, this was not true for more complex constructions (3P past and present and quell wh-object question) children with ASD-LI were significantly more impaired than children with SLI (70% of the ASD-LI group displayed some errors vs. 46% in the SLI group). On the other two conditions (relative clauses and argument clauses) the rate of children with ASD-LI making repetition errors was always 100%. Notably on object and subject relatives and on finite argument clauses all 14 children with ASD-LI showed repetition errors. For nonfinite argument clauses conditions, 70% of the ASD-LI group displayed some errors. On the other hand, the SLI group was significantly less impaired on all four conditions displaying a mean rate of children making errors that ranged from 19% on nonfinite argument clauses to 57% on object relatives.

These results suggest that the children with ASD-LI seemed to treat computational complexity differently than the children with SLI. While the SLI group performed better on less complex substructures than on more complex ones on all five conditions, the ASD-LI group displayed this difference exclusively on monoclausal sentences and on object wh-questions. In contrast, no difference in performance was spotted in clauses involving embedding and/or movement in this group. Moreover, the children with SLI seemed to be more prone to keep the target structure of the sentences in the wh-object, relative clauses and argument clauses conditions, displaying higher performance on Target structure than on Identical repetition. In contrast, the children with ASD-LI made significant errors both on less complex and more complex constructions on Identical repetition score, indicating that other mechanisms may be behind their performance. We will come back to this idea in section 7.3.3.4 (developmental trajectories) and in the discussion of these results (section 7.3.4).

We now move to the comparison between the ASD-LN and TD groups. Figure 31 shows the performance of these groups on the different substructures.

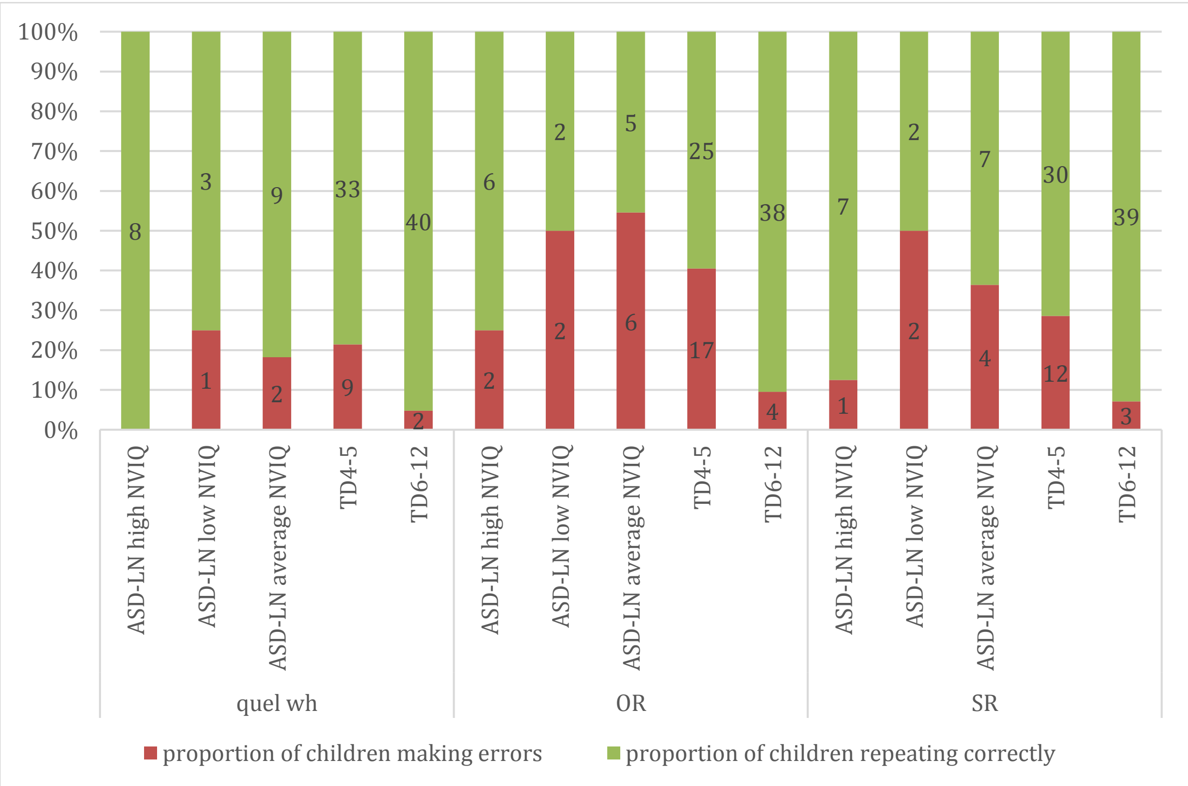
FIGURE 31. ASD-LN AND TD PERFORMANCE ON SUBSTRUCTURES IN THE SR TASK



The ASD-LN high NVIQ group did not differ from either TD group on any substructures. In contrast, the ASD-LN with low average NVIQ subgroup performed significantly below the TD6-12 on *quel* object wh-questions ($U(45) = 62.5$, $p = .002$), object relatives ($U(45) = 43.5$, $p = .005$) and subject relatives ($U(45) = 60$, $p < .001$), for both Identical repetition and Target structure. The ASD-LN with average NVIQ subgroup performed significantly below the TD6-12 group on object relatives ($U(50) = 152$, $p = .005$) and subject relatives ($U(50) = 140$, $p < .001$) on both Identical repetition and Target structure. None of the two ASD-LN groups differed from the TD4-5 group.

Looking at individual performance, we will concentrate only on the substructures that caused difficulties in the children with ASD-LN (*quel* object wh-questions and relative clauses). Figure 32 reports the proportion (and number) of children with 100% correct repetition (in green bars), versus those having less than 100% correct repetition (in red bars), on the three selected conditions.

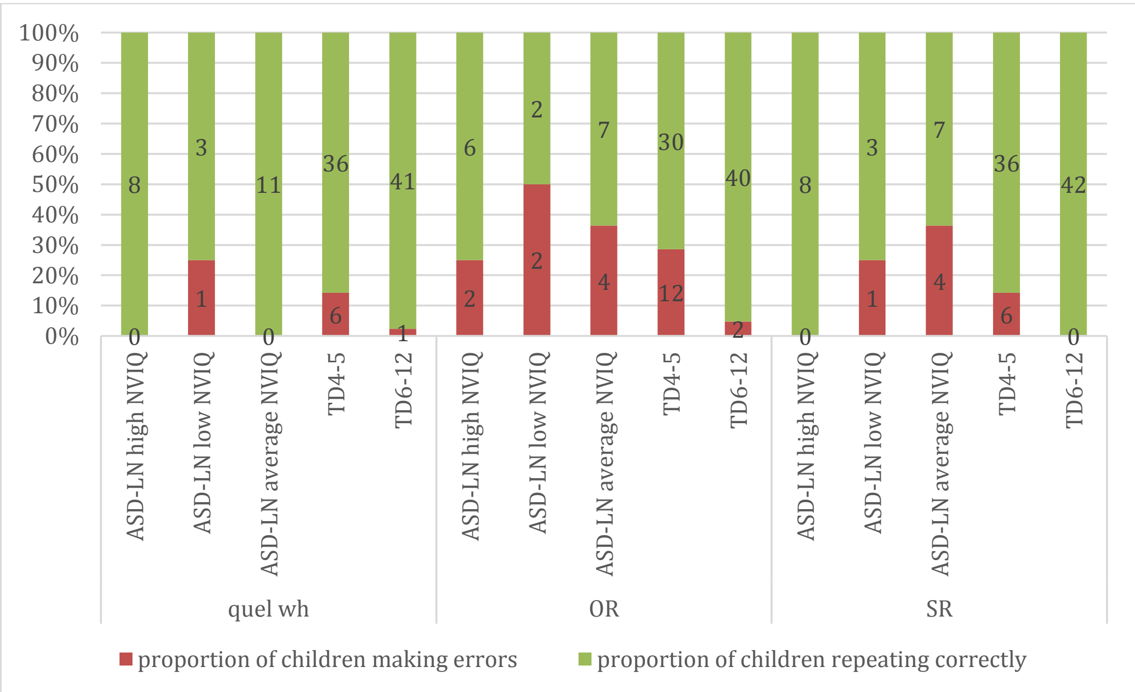
FIGURE 32. PROPORTION (AND NUMBER) OF CHILDREN IN THE ASD-LN AND TD GROUPS MAKING ERRORS AND CORRECTLY REPEATING SUBSTRUCTURES (IDENTICAL REPETITION SCORE)



The low performance on *quel* wh-object in the ASD-LN with low NVIQ and ASD-LN with average NVIQ groups was due to the low scores of one or two children who selectively performed low on these conditions and lowered the group scores. For relative clauses the number of children displaying errors was higher than for the wh-objet condition. In the ASD-LN with low NVIQ group 2/4 children displayed incorrect repetitions on both subject and object relatives. Similarly in the ASD-LN with average NVIQ group, half of the children (6/11) made errors on both object and subject relatives. Neither the ASD-LN with low NVIQ nor the ASD-LN with average NVIQ group differed from the TD4-5 group on any of the three substructures. In the TD4-5 group, in fact, roughly 40% of the children showed errors in repetitions on object relatives and 30% on subject relatives (percentages very similar to the ASD-LN groups).

Figure 33 shows individual results on Target structure scores. No difference was found between Identical repetition and Target structure scores for the ASD-LN groups, except for the *quel* wh-object condition in the ASD-LN with average NVIQ group. Both object and subject relatives still caused difficulties among the ASD-LN with low NVIQ and the ASD-LN with average NVIQ groups, indicating that these children struggled with these structures. In contrast, the children in the TD4-5 group showed significantly higher performance on Target structure scores, especially for subject relative clauses.

FIGURE 33. PROPORTION (AND NUMBER) OF CHILDREN IN THE ASD-LN AND TD GROUPS MAKING ERRORS AND CHILDREN REPEATING CORRECTLY ON SUBSTRUCTURES (TARGET STRUCTURE SCORE)



As mentioned above, comparing children with ASD-LN with children with TD might cause type 1 errors, since the number of participants in the ASD-LN groups is so small. Nonetheless, the Target structure scores show that at least on the object and subject relative conditions many children had difficulties, indicating that it could be interesting to focus an error analysis on these particular structures to examine in detail the types of errors made by the children. We return to this analysis in the next section (7.3.3.4.2).

In short, results on the ASD-LN groups suggested that except for the ASD-LN with high NVIQ group who performed in line with the age-matched TD6-12 group, both ASD-LN with average NVIQ and ASD-LN with low NVIQ children displayed some effect of computational complexity on relative clauses, even when compared with younger TD4-5 children.

7.3.3.3.2 Intragroup comparison

Two intragroup comparisons were run: the first one concentrated on statistical differences between Identical repetition scores and Target structure scores for each substructure in each group. This analysis aimed to see whether children performed differently on the two scores. While the three ASD-LN groups and the TD6-12 group did not differ between Identical repetition and Target structure scores, the ASD-LI, the SLI and the TD4-5 groups displayed some significant differences, always in favour of target structure scores. Table 34 reports differences showed by the three groups on each substructure.

TABLE 34. DIFFERENCES BETWEEN IDENTICAL REPETITION AND TARGET STRUCTURE SCORES, BY SUBSTRUCTURE

		3P Present	3S present	3P past	3S past	Quel	Qui	OR	SR	Finite	Nonfinite
ASD-LI	Z	-1.890	-2.719	-2.392	-2.724	-2.070	-1.342	-1.633	-1.841	-1.342	-2.836
	p	.059	.007**	.017*	.006**	.038*	.180	.102	.066	.180	.005**
SLI	Z	-1.897	-1.518	-1.511	-1.651	-2.271	-.687	-1.000	-2.121	-2.675	-3.666
	p	.058	.129	.131	.099	.023*	.492	.317	.034*	.007**	<.001***
TD4-5	Z	-2.000	-1.732	-2.414	-1.732	-1.732	-1.000	-3.162	-2.646	-4.823	-5.150
	p	.046*	.083	.016*	.083	.083	.317	.002**	.008**	<.001***	<.001***

Key: * p <.05; ** p < .01; *** p < .001

In second instance, we compared children's performance on less complex and more complex constructions both Identical repetition and target Structure scores, in order to see if the groups performed better on structures that involved less computational complexity than on structures involving more computational complexity. Results are reported in Table 35. The analysis indicated that on identical repetition, children with ASD-LI displayed more difficulty on more complex object wh-questions and relative clauses than on their less complex counterparts, while children with SLI seemed to display computational complexity effects on present tense SVO and object wh-questions. On Target structure scores, the ASD-LI group systematically showed difference of performance between less complex and more complex substructures on all five conditions, while for the SLI group differences were limited to present tense SVO, relative clauses and argument clauses. The three ASD-LN groups and the TD 6-12 group did not show any difference in performance between less complex and more complex substructures, while results for the TD4-5 group point to a difference on past tense SVO, object wh-questions and relative clauses on Identical repetition scores and on object wh-questions, relative clause and argument clauses on Target structure scores.

TABLE 35. COMPARISON BETWEEN PERFORMANCE ON LESS COMPLEX AND MORE COMPLEX SUBSTRUCTURES FOR EACH GROUP

Groups		Identical repetition					Target structures				
		Present tense 3P/3S	Past Tense 3P/3S	Wh-object quel/qui	Relative clauses OR/SR	Argument clauses finite/nonfinite	Present tense 3P/3S	Past Tense 3P/3S	Wh-object quel/qui	Relative clauses OR/SR	Argument clauses finite/nonfinite
ASD-LN with high NVIQ	Z	-1.414	-1.414	-1.414	-.816	-.378	-1.000	.000	-1.000	-1.414	-1.000
	p	.157	.157	.157	.414	.705	.317	1.000	.317	.157	.317
ASD-LN with low NVIQ	Z	.000	-1.414	-.447	-.447	-1.134	.000	.000	-1.000	-.816	.000
	p	1.000	.157	.655	.655	.257	1.000	1.000	.317	.414	1.000
ASD-LN with average NVIQ	Z	-1.414	-1.342	-.816	-1.732	-.302	-1.000	-1.732	.000	-1.000	-1.342
	p	.157	.180	.414	.083	.763	.317	.083	1.000	.317	.180
TD4-5	Z	-1.513	-2.624	-2.233	-2.388	-1.091	-1.667	-1.933	-1.994	-2.244	-2.271
	p	.130	.009	.026	.017	.275	.096	.053	.046	.025	.023
TD6-12	Z	.000	-1.000	-1.342	-.632	-.836	-1.000	.000	-1.000	-1.342	-1.890
	p	1.000	.317	.180	.527	.403	.317	1.000	.317	.180	.059
ASD-LI with (low) average NVIQ	Z	-.531	-1.276	-2.980	-2.126	-.577	-2.579	-2.558	-2.724	-2.598	-2.558
	p	.595	.202	.003	.033	.564	.010	.011	.006	.009	.011
SLI	Z	-2.652	-1.592	-2.275	-1.642	-1.795	-2.165	-1.903	-1.026	-2.019	-3.170
	p	.008	.111	.023	.101	.073	.030	.057	.305	.043	.002

Note: grey-cells indicates significant differences

Results on intragroup comparisons suggested that it should be more correct to consider both Identical repetition and Target structure scores at least for ASD-LI and SLI groups, for not falling into underestimation or overestimation of their linguistic abilities. Nonetheless, further analysis investigating potential qualitative similarities and differences between the ASD-LI and SLI groups, with regarding to errors on substructures that has been pinpointed as different, is needed.

7.3.3.4 Error analysis

In this section we will investigate similarities and differences between children with ASD-LI and children with SLI and between the two ASD-LN groups (ASD-LN with low NVIQ and ASD-LN with average NVIQ) who displayed significant differences with the TD groups in a qualitative error analysis perspective. This kind of analysis is fundamental in providing a detailed picture of the types of errors made in each group. We will use this information in our delineation of phenotypical languages profiles of children with ASD. Firstly we will concentrate on performance in the ASD-LI and SLI groups and then we will move to the ASD-LN and TD groups.

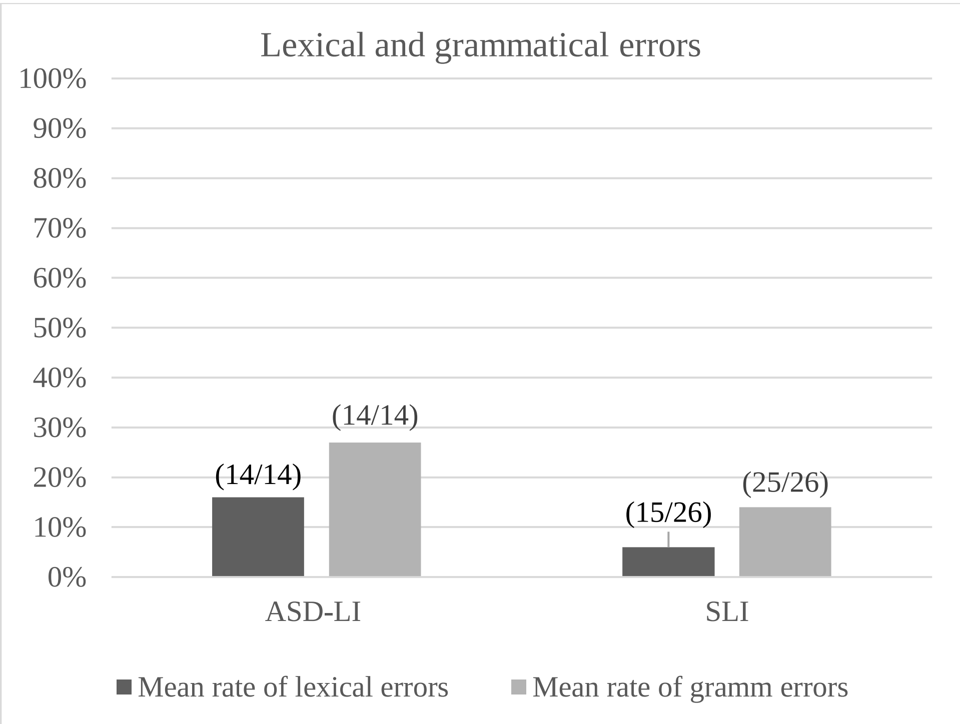
7.3.3.4.1 ASD-LI and SLI

Our conclusion regarding group global results and individual global results suggested that the children with ASD-LI and with SLI seemed to differ, at some level, for severity of impairment. The children in the ASD-LI group seemed to be more impaired than the children with SLI, repeating substructures at lower rates and being more affected as a group (more children with very low performance on both identical repetition and target structure scores). The question that then arises is what is behind this ASD-LI/SLI difference. As part of answering this question, we would like to know whether the two groups of children differed on error typology, whether they differed on mean rates of errors (more errors in the ASD-LI group than in the SLI group), or both.

To answer this question we will first report on error typology, describing main error categories (in terms of the types of words/morphemes that were affected by an error, lexical or grammatical) over the total errors made in each of the ASD-LI and the SLI groups. Then we will focus on the types of errors that were specific to each substructure, analyzing preferentially the substructures on which the two groups performed differently in the general analysis. For each of these error analyses, we will report the number of children concerned for each error type.

As a group, the 14 children with ASD-LI made 835 errors in the SR task, while the 26 children with SLI made 732 errors. Among the total errors a first distinction can be made between "lexical errors" and "grammatical errors". "Lexical errors" were substitutions and omission of lexical items, while "grammatical errors" were substitutions and omissions of grammatical morphemes (words or inflectional affixes). In the ASD-LI group, 245 out of the 1498 total lexical items that these 14 children had to repeat were omitted or substituted; for grammatical items the number of omissions and substitutions was 386 of 1400. In the SLI group, 189 out of the total 2782 lexical items were omitted or substituted, while for lexical items the number of omissions and substitutions was 383 out of a total of 2600. Mean error rates over the total number of targeted lexical and grammatical items are reported in Figure 34.

FIGURE 34. MEAN RATE OF LEXICAL ERRORS AND GRAMMATICAL ERRORS FOR THE ASD-LI AND SLI GROUPS; N OF PARTICIPANTS INVOLVED IN PARENTHESIS



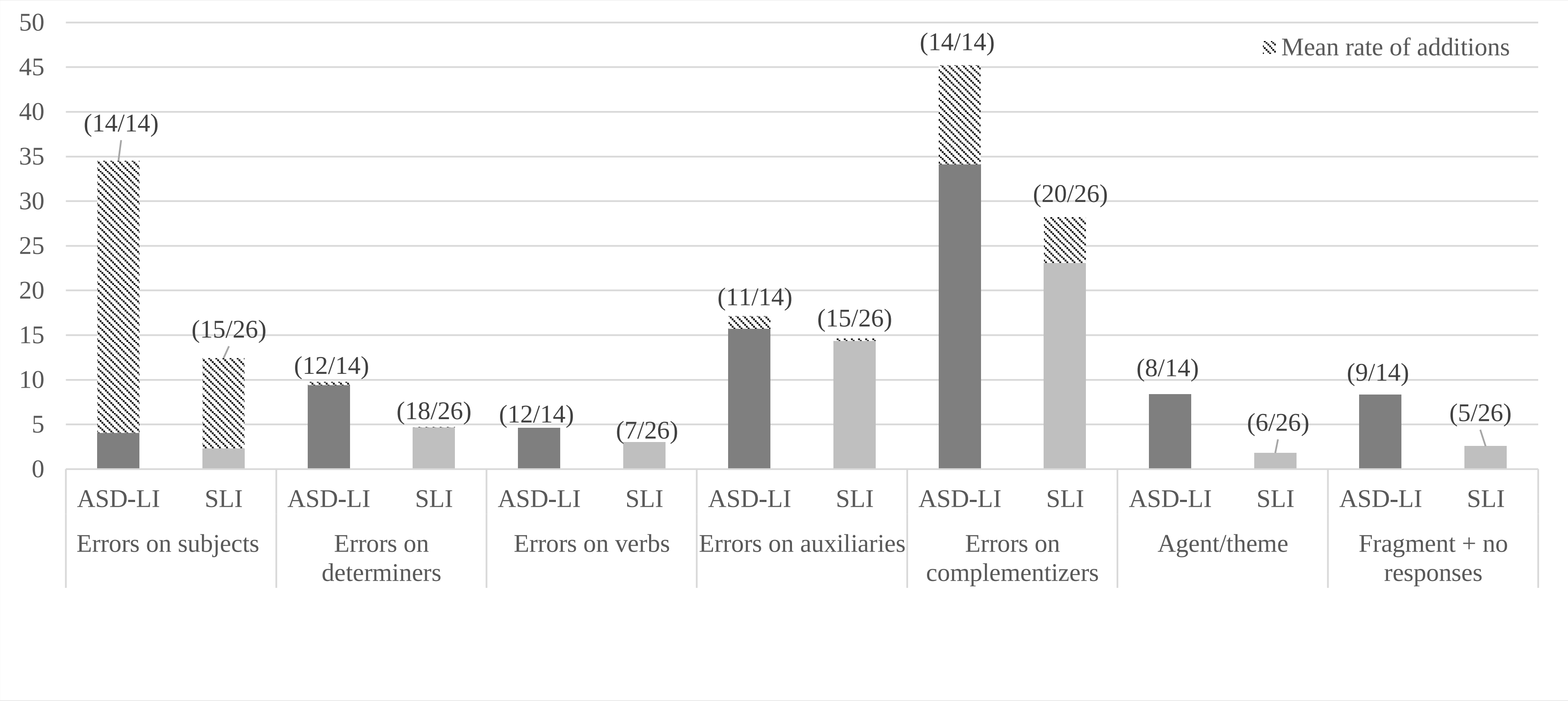
Significant differences were found between the two groups for both the total number of lexical errors ($U(39) = 61.5$, $p = .001$) and grammatical errors ($U(39) = 59$, $p < .001$). Not only was the number of errors systematically higher in the ASD-LI group than in the SLI group, but the proportion of children making errors in the ASD-LI group was much higher than in the SLI group; while the proportion of children in the ASD-LI group was 100%, in the SLI group the mean rate of children making errors was between 60% (on lexical errors) and 88% (grammatical errors). Again, the discussion behind the possible reasons of these differences will be addressed in section 7.4.4.5 and in the discussion of the present work. Besides errors involving substitution or omission of lexical and grammatical morphemes, children also added morphemes that were not in the sentence they heard. In the ASD-LI group, there were additions of 5 lexical items (0.33%) and 199 (14%) grammatical items, while the SLI group added 41 lexical items (1.4%) and 154 (6%) grammatical items.

Moving to a more specific error analysis, we calculated seven scores that summed up similar types of errors across all sentences in the task.

(1) Errors on subjects: this score represents the sum of errors affecting subject (clitic or DP) substitutions and omissions.
(2) Errors on determiners: this score represents the sum of determiner omissions and substitutions, and gender (feminine vs masculine) and number (singular vs plural) errors.
(3) Errors on verbs: this score includes errors on verb number agreement (singular vs plural) and tense-related errors (finite vs. nonfinite).
(4) Errors on auxiliaries: this score sums up substitution errors (be vs have), auxiliary omissions and number agreement errors (singular auxiliary for a plural form).
(5) Errors on complementizers: this score accounts for substitutions and omissions of complementizers (this score concerns relative clauses and argument clauses only).
(6) Agent/theme reversal: this score consists of errors that involved an inversion of agent/theme roles on animate DPs.
(7) Fragment + No responses: this score consists of all productions where children produced only one word (or phrase) and when they did not repeat the sentence at all.

Figure 35 shows the mean error rate for each error type calculated over the total number of occurrences for subject, determiner, verb, auxiliary and complementizer errors, and over the total number of sentences containing two animate DPs for "Agent/theme reversal errors" and over the total number of sentences on "Fragment + no responses errors" in each group. The striped bars show addition errors for both groups. For subject addition, this score reflected the addition of a subject clitic linked to the subject of the sentence as exemplified in (13b).

FIGURE 35. MEAN RATE OF ERROR TYPES OVER THE TOTAL OCCURRENCES INCLUDING ADDITIONS (N OF PARTICIPANTS INVOLVED)



Concerning omission errors and subtitution errors, significant differences were found between the ASD-LI and SLI groups on “Errors on determiners” ($U(39) = 97$, $p = .015$), “Errors on complementizers” ($U(39) = 107.5$, $p = .033$), “Agent/Theme reversal errors” ($U(39) = 107$, $p = .012$) and “Fragment + no responses” ($U(39) = 99.5$, $p = .006$). Including addition errors the significant differences also extended to “Errors on subjects” ($U(39) = 59$, $p < .001$). Again, in line with what was indicated for global results, while (almost) all children in the ASD-LI group made errors on subjects, determiners, verbs, auxiliaries and complementizers, the rate of children involved in making errors among the SLI group was never higher than 77% (errors on complementizers). Taking the example of “Errors on determiners” the total number of errors was very similar in the two groups, but on average children with ASD-LI made 6.25 errors each, while children with SLI made 3.8 errors each. This was partially explained by the fact that in the ASD-LI group some children made much more errors than the others. Concerning “Agent/theme reversal” and “Fragment + no responses”, roughly 57 % of the children with ASD-LI made this type of errors, while for SLI such errors were found in only 20% of the group.

Concentrating next on individual substructures, we sought to identify which errors were the most frequent in each group for each condition. We saw in section 7.3.3.3 that children with ASD-LI differed from children with SLI on all substructures (except for present tense 3P and *qui* object wh-questions) on Identical repetition scores. Significant differences were circumscribed to relative clauses and argument clauses when we took into account Target structure scores. Our goal in identifying which types of errors were the most frequent in substructures was to seek an explanation for why the significant ASD-LI/SLI differences disappeared when target structure scores were considered, and furthermore to see which errors (and how many) might be more related to computational complexity (the underlying structure of the sentence). We will give one representative example for each type of error; sentences that are ungrammatical in French will be indicated with an asterisk (*).

3S present tense

Two main types of repetition errors were spotted in 3S present tense substructures: omission of determiners, as in (13a), and subject clitic addition, as in (13b).

(13) a. Input : *La* *maman lit* *une* *histoire*

The.fem.sing mum reads.3sing a.fem.sing story

(Structure: SVO present tense 3S)

b. Child repetition: Maman lit une histoire

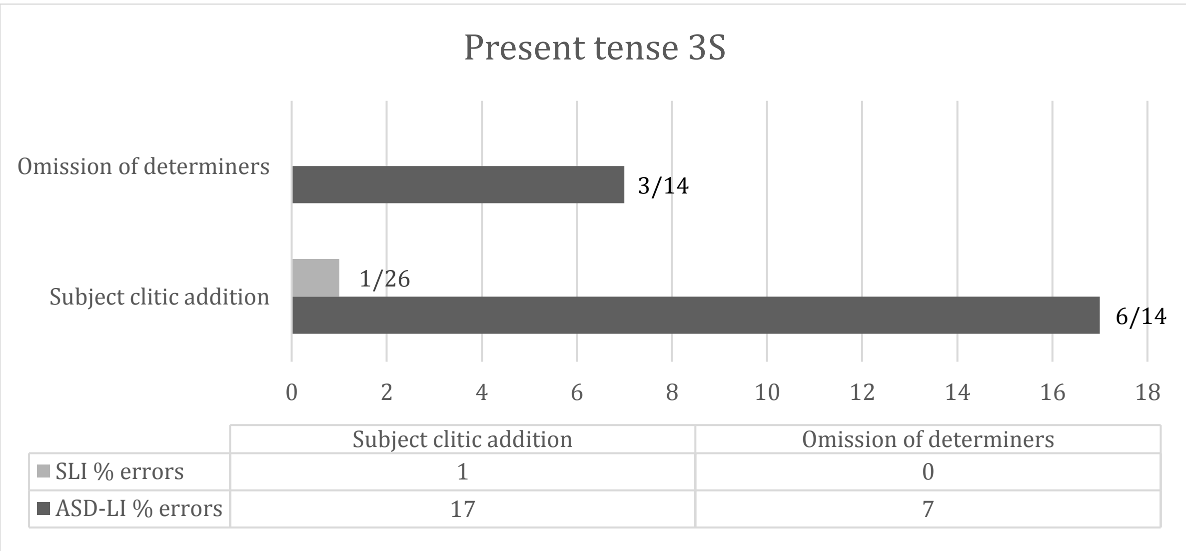
'Mum reads a story'

c. Child repetition : La maman elle lit une histoire

'The mum she reads a story'

Figure 36 shows the error rate for all errors on present tense 3S substructures, in each group.

FIGURE 36. ERROR TYPES FOR 3S PRESENT TENSE



These two main error types were found only in the ASD-LI group (except for one child in the SLI group). These qualitative results explain why the significant difference between the ASD-LI and SLI groups obtained on identical repetition vanished when the two groups' performance was compared on Target structure scores. Since both "omission of determiners" and "subject clitic addition" do not alter the target structure of the sentence, these errors were not counted in the

Target structure scores. When the Target structure score is taken into account, then, the two groups performed alike, as shown in Figure 28.

3P and 3S past tense

Three main types of errors were found for the past tense 3P condition: subject clitic addition, auxiliary omission, as in (14a), and errors of number agreement on auxiliaries (singular for plural) as in (14c):

(14) a. Input : *Les parents ont rangé les jouets*

The.pl parents have.3pl put+away the.pl toys

(Structure: SVO past tense 3P)

b. Child repetition: * Les parents rangé les joutes

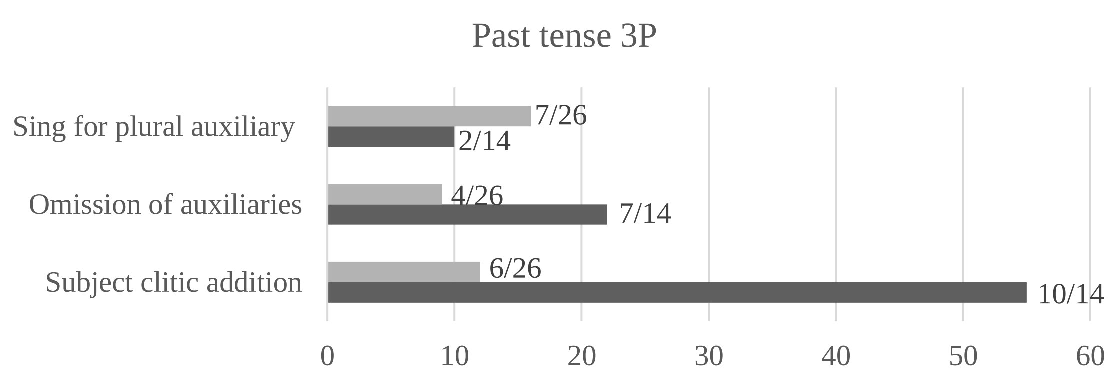
‘The parents put away the toys’

c. Child repetition : * Les parents a rangé les jouets

‘The parents has put away the toys’

Figure 37 shows the mean rate of errors on the total errors on past tense 3P substructures for each group.

FIGURE 37. ERROR TYPES FOR 3P PAST TENSE



	Subject clitic addition	Omission of auxiliaries	Sing for plural auxiliary
SLI % errors	12	9	16
ASD-LI % errors	55	22	10

For the past tense 3S condition children made mainly three types of errors: subject clitic addition, omission of auxiliaries and gender errors on determiners, as shown in (15).

(15) a. Input : *Le singe* *a* *pris la* *banane*

The.masc.sing monkey have.3sing taken the.fem.sing banana

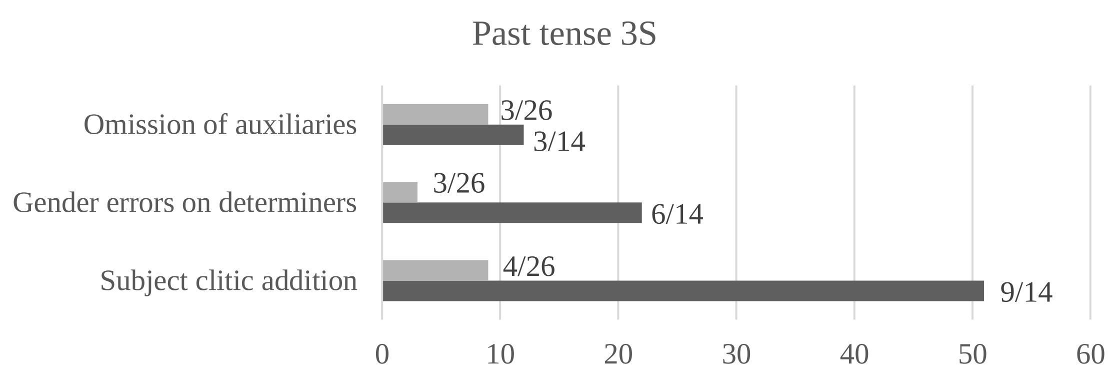
(Structure: SVO past tense 3S)

b. Child repetition: * Le singe il a pris le banane

'The monkey he has taken the.masc.sing. banana'

Figure 38 shows the mean rate of errors on the total errors on past tense 3S substructures for each group.

FIGURE 38. ERROR TYPES FOR 3S PAST TENSE



	Subject clitic addition	Gender errors on determiners	Omission of auxiliaries
SLI % errors	9	3	9
ASD-LI % errors	51	22	12

For both these conditions the most frequent error in the ASD-LI group was subject clitic addition. As expected, for Target structure scores the difference in performance between the two groups was no longer significant (Figure 28). Regarding the other errors spotted in each of the groups it is interesting to notice that, although the number of errors wass not very high and only concerned a few children in each group, there were some differences. While in the ASD-LI group some children omitted the auxiliary in their repetition of past tense 3P, children with SLI tended to use a singular auxiliary instead of a plural one. We note that this difference between omission and substitution of grammatical morphemes was found for other substructures (object relatives and finite argument clauses) as well; we return to this result in the discussion (section 7.3.4). On the past tense 3S condition almost half of the ASD-LI group produced gender errors on determiners, an error rarely found in the SLI participants.

Quel object wh-question

Six main error types were found for *quel* object wh-questions: subject clitic addition, determiner omission, complementizer addition, as shown in (16), subject wh-questions for object wh-questions, as in (17), agent/theme reversal, as in (18), and fragment + no responses.

(16) a. Input : *Quel garcon le monsieur dessine?*

Which.masc.sing boy the.masc.sing man draws.3sing?

(Structure: *quel* object wh-questions)

b. Child repetition: *Quel garçon que le monsieur dessine ?

'Which boy that does the man draw?'

(17) a. Input : *Quel enfant la maîtresse punit?*

Which.masc.sing boy the.fem.sing teacher punishes.3sing?

(Structure: *quel* object wh-questions)

b. Child repetition: *Quel enfant qui punit la *maîtresse*?

'Which boy who does punish the teacher?'

(18) a. Input : *Quel garçon le papy connait?*

Which.masc.sing boy the.masc.sing grandpa knows.3sing?

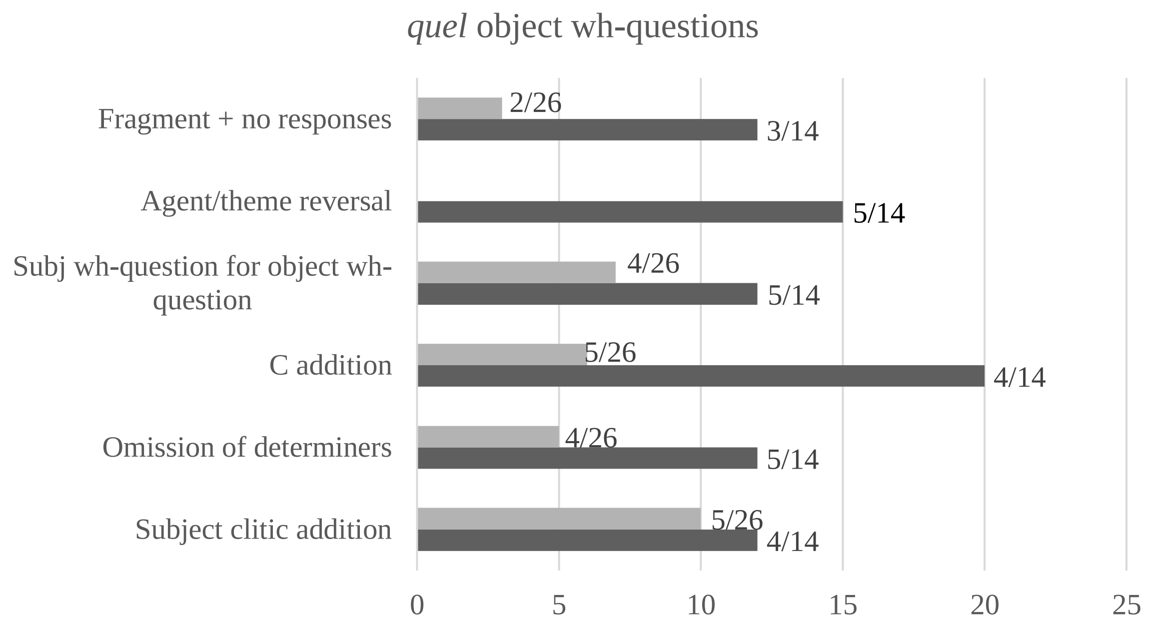
(Structure: *quel* object wh-questions)

b. Child repetition: *Quel papy que le garçon connait?

'Which grandpa that does the boy know?'

Figure 39 shows the mean rate of errors on the total errors on *quel* object wh-questions for each group.

FIGURE 39. ERROR TYPES FOR QUEL OBJECT WH-QUESTIONS



	Subject clitic addition	Omission of determiners	C addition	Subj wh-question for object wh-question	Agent/theme reversal	Fragment + no responses
SLI % errors	10	5	6	7	0	3
ASD-LI % errors	12	12	20	12	15	12

Children in the ASD-LI group and the SLI group showed marked differences in error typology, such as addition of complementizers (C addition) which was higher in the ASD-LI group, and agent/theme reversal which was absent in children with SLI. Moreover, we can see that in both groups the number of children producing errors was very similar: 1/3 of the children in the ASD-LI group and 1/5 of the children in the SLI group. The relatively low rate of children producing errors on these constructions and the relative heterogeneity of error typology did not allow for strong conclusions.

Up until now our qualitative error analysis focused on substructures for which performance in the ASD-LI and SLI groups differed only on Identical repetition score. We now move to qualitative error analysis of sentences including embedded clauses, which showed significant differences on both Identical repetition and Target structure scores.

Object and subject relatives

TABLE 36. MEAN RATE OF CORRECT REPETITION AND ERRORS ON IDENTICAL REPETITION AND TARGET STRUCTURE SCORES FOR THE ASD-LI AND SLI GROUPS

		ASD-LI		**SLI**	
		Identical repetition score	Target structure score	Identical repetition score	Target structure score
Object relatives	% Correct	5%	15%	53%	58%
	% Errors	95%	85%	47%	42%
Subject relatives	% Correct	24%	41%	64%	72%
	% Errors	76%	59%	36%	28%

Table 36 there shows the mean rate of correct repetition and errors on both Identical repetition and Target structure scores for ASD-LI and SLI groups on relative clauses. We can see that children with ASD-LI obtained rates of errors roughly twice as high as children with SLI. Which were the reasons behind this difference? Nine main types of errors were founded for object relatives, three of which were found only for identical repetition: subject clitic addition, gender errors on determiners, and tense substitutions, as in (19b). Six were found for both identical repetition and target structure scores: omission of the complementizer, as in (19c), substitution of the complementizer, as in (19d), production of a monoclausal, as in (19e), use of an subject relative instead of an object relative, as in (19f), agent/theme reversal and fragment + no responses.

(19) a. Input : *Tu as vu la vache que le chat a griffé*

You saw the.fem.sing cow that the.masc.sing cat scratched

(Structure: object relatives)

b. Child repetition: Tu vois la vache que le chat a griffé

'You see the cow that the cat scratched'

c. Child repetition: * Tu as vu la vache le chat a griffé

'You saw the cow the cat scratched'

d. Child repetition: * Tu as vu la vache le chat qui a griffé

'You saw the cow the cat who scratched'

e. Child repetition: La vache il a griffé le chat

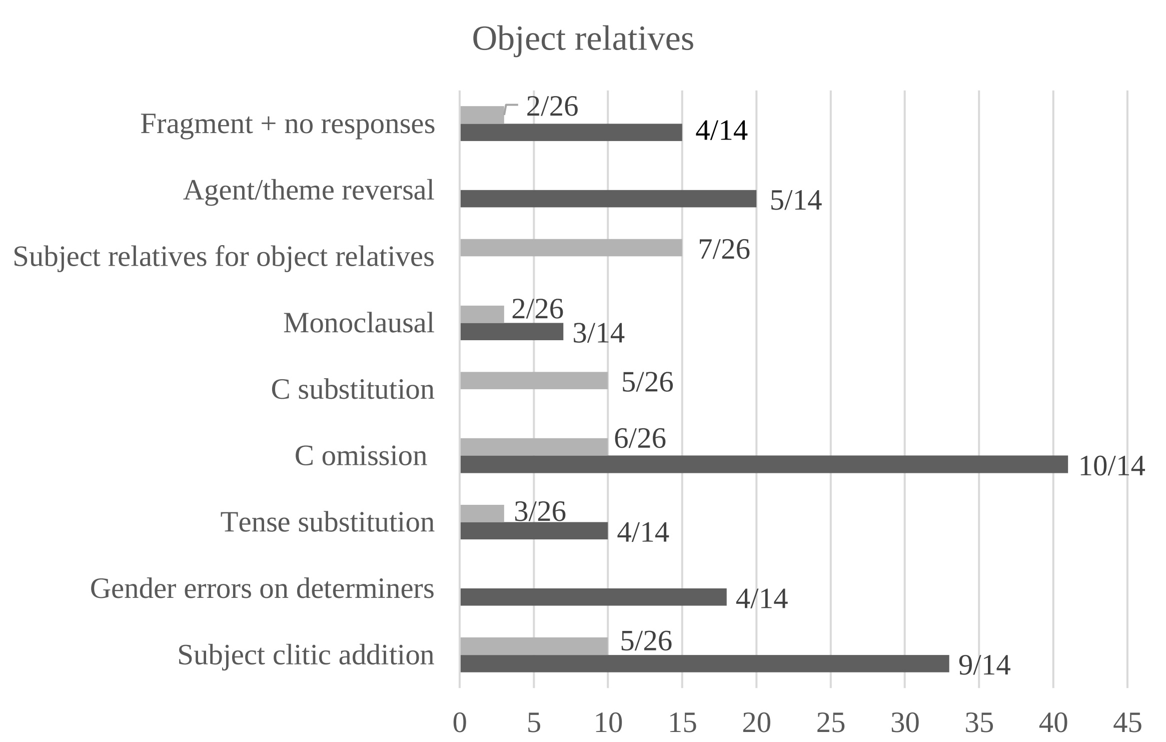
'The cow he scratched the cat'

f. Child repetition: Tu as vu la vache qui a griffé le chat

'You saw the cow who scratched the cat'

Figure 40 shows the mean rate of errors on the total errors on object relative substructures, for each group.

FIGURE 40. ERROR TYPES FOR OBJECT RELATIVES



	Subject clitic addition	Gender errors on determiners	Tense substitution	C omission	C substitution	Monoclausal	Subject relatives for object relatives	Agent/theme reversal	Fragment + no responses
SLI % errors	10	0	3	10	10	3	15	0	3
ASD-LI % errors	33	18	10	41	0	7	0	20	15

Qualitative error analysis on object relatives showed some differences in error typology between the ASD-LI and SLI groups. Concerning errors found only for Identical repetition scores, the children with ASD-LI showed a higher rate of subject clitic addition than the children with SLI (twice higher, 33% vs. 10% and three times as frequent in terms of children involved, 64% vs. 20%). Moreover they produced gender errors on determiners, an error that was not found in the SLI group.

Concerning errors linked to the structure of the object relative clause, the children with ASD-LI made four types of errors: omission of the complementizer (41%), production of a monoclausal (7%), agent/theme reversal (20%) and fragment + no responses (15%). Children with SLI seemed to be more prone to substitute the complementizer (*que* > *qui*, 10%) and change the structure from object relatives to subject relatives (15%), along with a few monoclausal errors (3%) and fragment + no responses (3%). If we put together all the errors on Target structure in each group we obtain 83% vs 41% of target errors respectively, which roughly corresponds to the mean rate of errors on target structure reported in Table 36. No child with ASD-LI produced subject relatives to replace object relatives or C substitutions. Similarly to what we saw before for the past tense condition, the children with ASD-LI made on gender errors determiners and, like on *quel* object wh-questions, they produced agent/theme reversals. No child with SLI made these kinds of errors. For tense substitutions, monoclausal and fragment + no responses, a slightly higher mean rate of errors was found in the ASD-LI group.

Concerning subject relative clauses, eight main errors were found. Three were found only for Identical repetition: determiner omission, auxiliary omission and subject clitic addition related to a monoclausal, as in (20b), while five were found for both Identical repetition and Target structures scores: C omission, production of a monoclausal, juxtaposition, as in (20c), argument clauses for subject relative, as in (20d), and fragment + no responses.

(20) a. Input : *Tu vois le garçon qui a dessiné la mamie*

You see the.masc.sing boy who drew the.fem.sing gradma

(Structure: subject relatives)

b. Child repetition: Le garçon il a dessiné la mamie

'The boy he drew the grandma'

c. Child repetition: * Tu vois le garçon il a dessiné la mamie

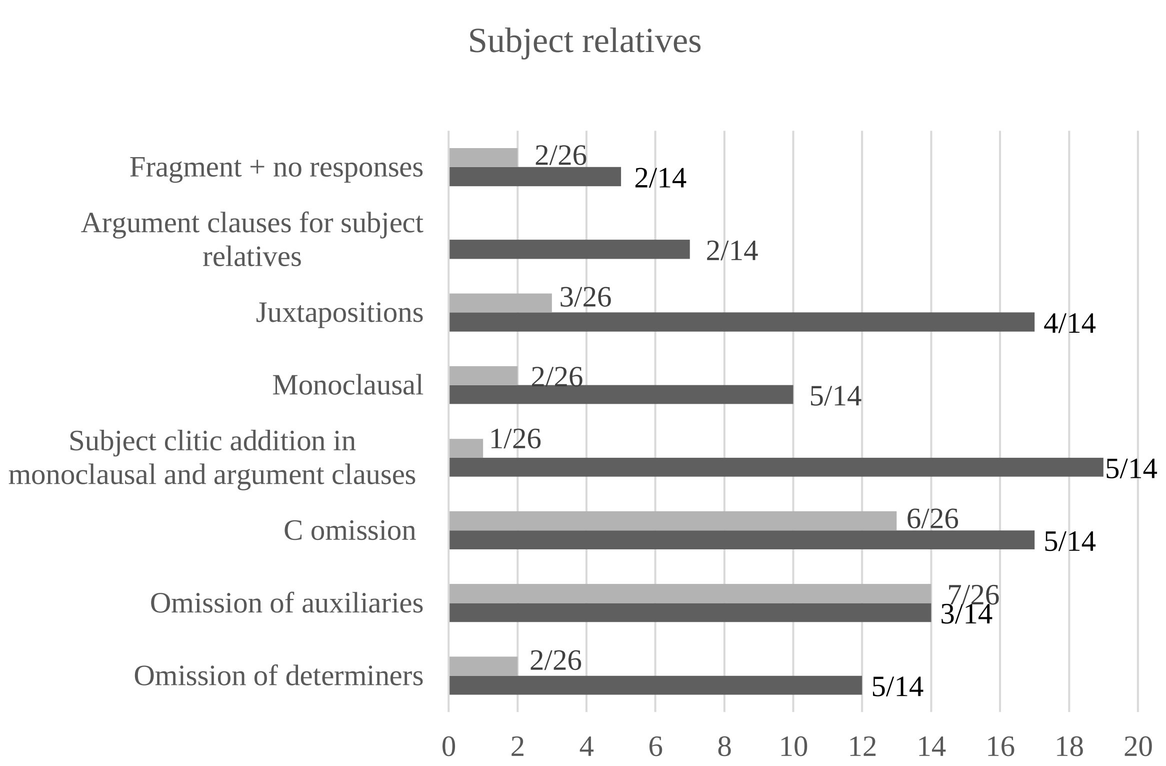
'You see the boy he drew the grandma'

d. Child repetition: Tu vois que le garçon il a dessiné la mamie

'You see that the boy he drew the grandma'

Figure 41 shows the mean rate of errors out of the total number of errors on subject relative substructures for each group.

FIGURE 41. ERROR TYPES FOR SUBJECT RELATIVES



	Omission of determiners	Omission of auxiliaries	C omission	Subject clitic addition in monoclausal and argument clauses	Monoclausal	Juxtapositions	Argument clauses for subject relatives	Fragment + no responses
SLI % errors	2	14	13	1	2	3	0	2
ASD-LI % errors	12	14	17	19	10	17	7	5

The ASD-LI and SLI groups showed some differences on error typology on subject relatives. Concerning errors that were found only for Identical repetition, differences were found on determiner omission which was more frequent in the children with ASD-LI than in the children with SLI, while on auxiliary omission the two groups largely resembled each other. Concerning errors more related to the structure of subject relatives five main errors were made by the children with ASD-LI: omission of the complementizer (17%), production of a monoclausal (10%) – sometimes with the addition of a subject clitic pronoun, juxtaposition of two sentences and thus elimination of the dependency between the main clause and the embedded clause (17%), an argument clause instead of an object relative (7%) and “fragment no response” (5%).

Except for argument clauses used instead of object relatives, the children with SLI made the same types of errors as the ASD-LI children, but at lower rates. Again, putting together all the found on Target structure in each groups we obtain 56% and 21% of target errors respectively, which roughly corresponds to the mean rate of errors on target structure reported in Table 36.

Finite and Nonfinite argument clauses

TABLE 37. MEAN RATE OF CORRECT REPETITION AND ERRORS ON IDENTICAL REPETITION AND TARGET STRUCTURE SCORES FOR ASD-LI AND SLI GROUPS

		ASD-LI		**SLI**	
		Identical repetition score	Target structures score	Identical repetition score	Target structures score
Finite arg. clause	% Correct	5%	15%	32%	54%
	% Errors	95%	85%	68%	46%
Nonfinite arg. clause	% Correct	7%	60%	47%	88%
	% Errors	93%	40%	53%	12%

Table 37 carries the same function as Table 36. Nine main errors were made on finite argument clauses: four were found for Identical repetition: subject clitic addition, omission of determiners, tense substitution, and omission of auxiliaries, while five were found both for Identical repetition and Target structure: C omission, production of a monoclausal, preposition addition, as in (21), agent/theme reversal, and fragment + no responses.

(21) a. Input : *Le garçon a dit que la maman a lu le livre*

The.masc.sing boy has said that the.fem.sing mum has read the.masc.sing book

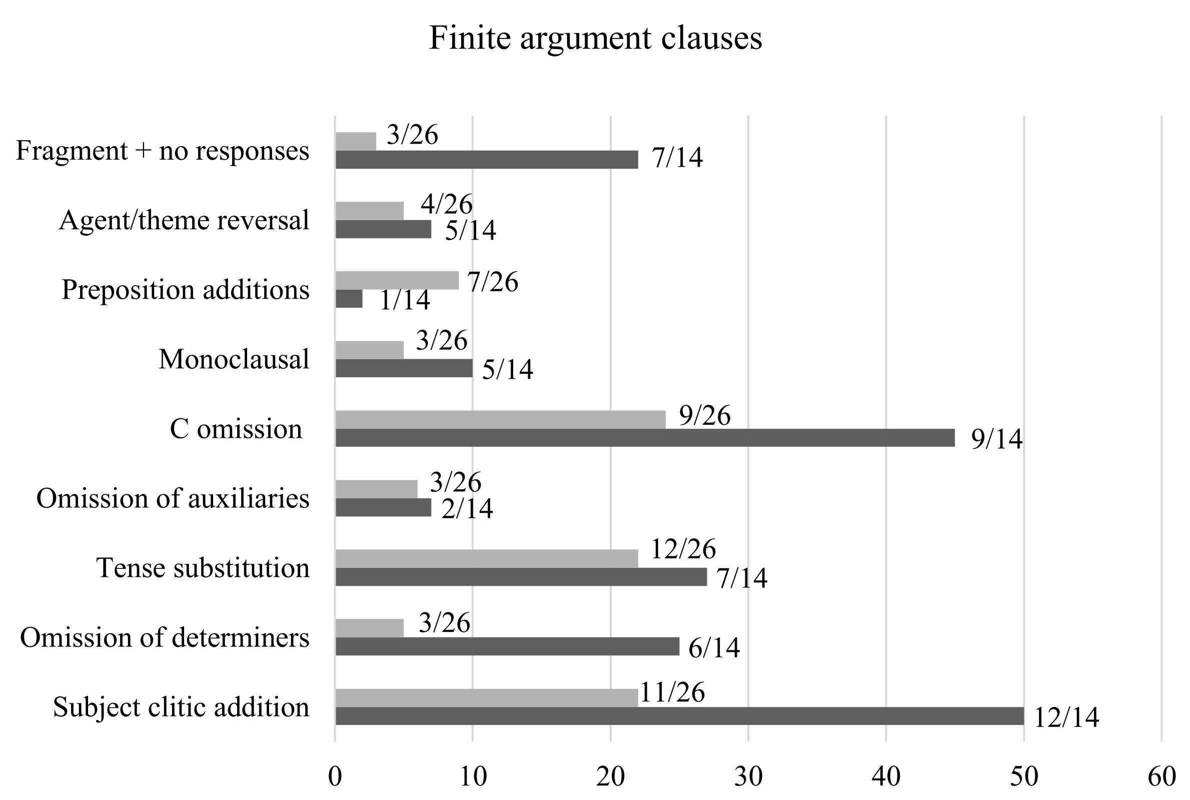
(Structure: finite argument clauses)

b. Child repetition: * Le garçon a dit à la maman qui a lu le livre

'The boy has said to the mum who has read the book'

Figure 42 shows the mean rate of errors out of the total number of errors on finite argument substructures for each group.

FIGURE 42. ERROR TYPES FOR FINITE ARGUMENT CLAUSES



	Subject clitic addition	Omission of determiners	Tense substitution	Omission of auxiliaries	C omission	Monoclausal	Preposition additions	Agent/theme reversal	Fragment + no responses
SLI % errors	22	5	22	6	24	5	9	5	3
ASD-LI % errors	50	25	27	7	45	10	2	7	22

Along with subject clitic addition, which was twice as numerous (50% of errors in ASD-LI vs. 22% of errors in SLI group) and twice as frequent (85% of ASD-LI group vs. 42% of SLI group) in the children with ASD-LI than in the children with SLI, the two groups differed on omission of determiners (which was much more frequent in the ASD-LI group). Similar results were found for tense substitution and auxiliary omission. Concerning the embedded part of the sentence, the children with ASD-LI produced four main errors: C omission (45%), produciton of a monoclausal (10%), agent/theme reversal (7%) and fragment + no responses (22%). The children with SLI poduced more C omissions (24%), some monoclausals (5%), some agent/theme reversals (5%) and a few fragment + no responses (3%). Roughly 25% of the SLI group changed the finite argument clause into a declarative sentence via adding a preposition (9%), an error rarely found in the children with ASD-LI. Adding all errors on the Target structure score in each groups, we obtain 86% and 48% of target errors respectively, which roughly corresponds to the mean rate of errors on target structure reported in Table 37.

Finally for nonfinite argument clauses five errors were spotted: three related only to Identical repetition scores: subject clitic addition, definite determiners for indefinite determiners, as in (22b), and tense substitution, and three that were found both for Identical repetition and target structure scores: produciton of a monoclausal, main verb omission, as in (22c), and fragment + no responses.

(22) a. Input : *La maman sait très bien dessiner des lapins*

The.fem.sing mum can.3sing very well draw.inf some.pl rabbits

(Structure: nonfinite argument clauses)

b. Child repetition: La maman sait très bien dessiner les lapins

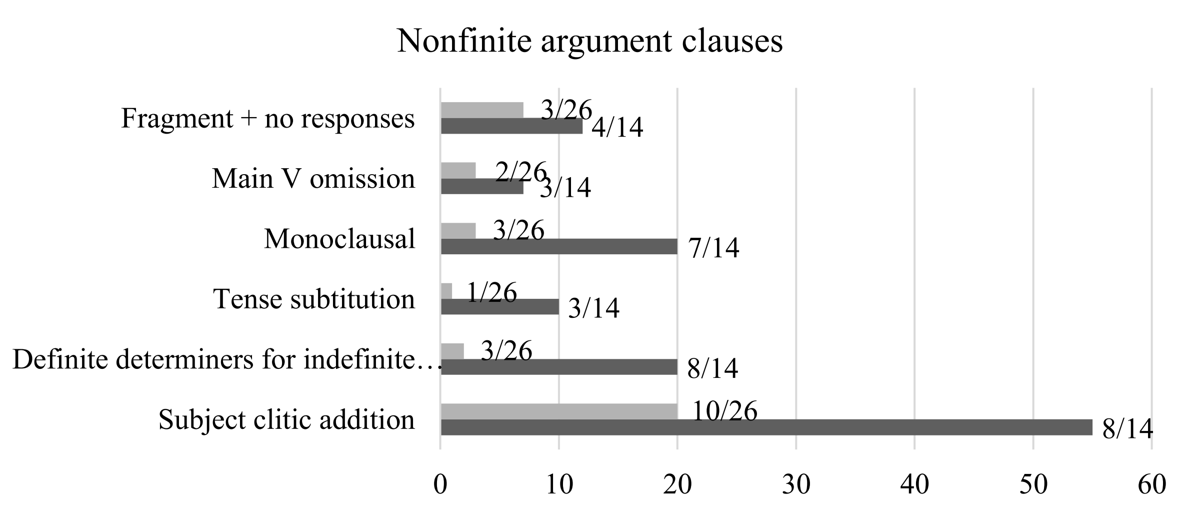
'The mum can very well draw the rabbits'

c. Child repetition: La maman très bien dessiner les lapins

'The mum very well draw the rabbits'

Figure 43 shows the mean rate of errors out of the total number of errors on nonfinite argument substructures for each group.

FIGURE 43. ERROR TYPES FOR NONFINITE ARGUMENT CLAUSES



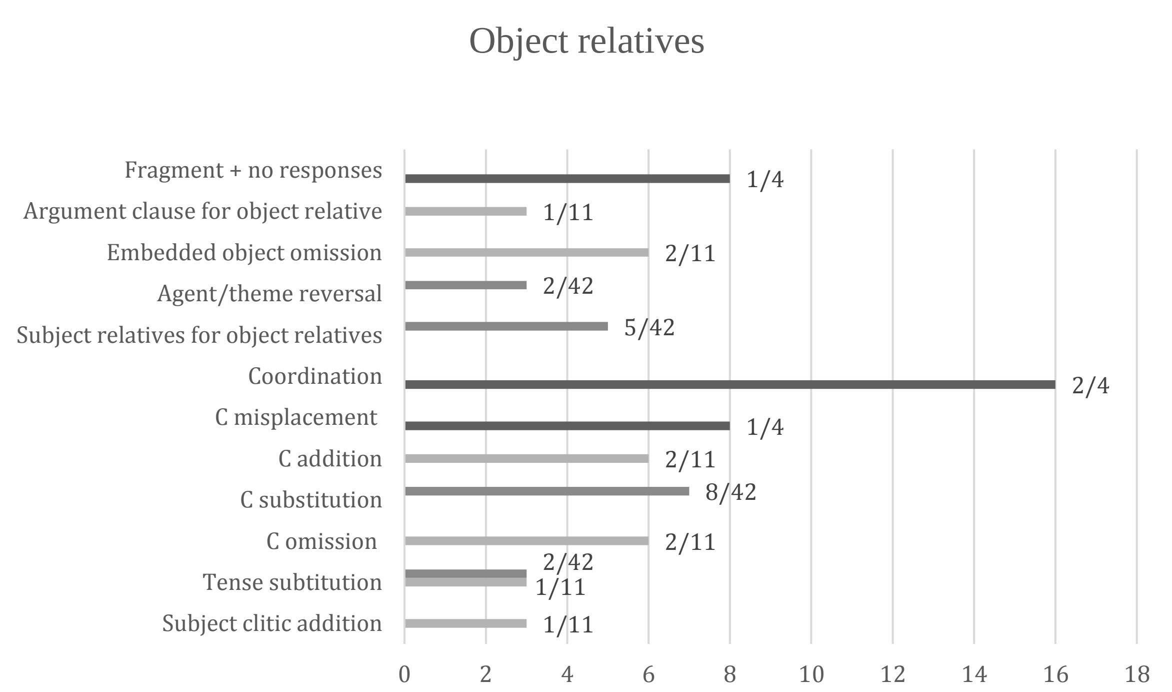
	Subject clitic addition	Definite determiners for indefinite determiners	Tense subtitution	Monoclausa l	Main V omission	Fragment + no responses
SLI % errors	20	2	1	3	3	7
ASD-LI % errors	55	20	10	20	7	12

For nonfinite argument clauses the two groups differed on subject clitic addition errors (twice as high and frequent in the ASD-LI group) and determiners errors (definite for indefinite, always more frequent among the children with ASD-LI). However, for determiners errors, we note that the exact same error (22b) was made by the children in the ASD-LI and SLI groups. Concerning the structure of the nonfinite clause, the children with ASD-LI produced more monoclausal errors (20%) than the children with SLI (3%) and while in the SLI group only three children were involved, in the ASD-LI group 50% of the children produced at least one monoclausal error. Adding together all the errors on Target structure we obtain 39% for ASD-LI group and 13% for SLI group (see Table 37).

7.3.3.4.2. ASD-LN and TD

Concerning the ASD-LN profiles, while the ASD-LN with high NVIQ group did not show any differences with the TD groups, the ASD-LN with low NVIQ group and the ASD-LN with average NVIQ group performed in line with younger TD4-5 children. All three groups displayed lower scores than the TD6-12 group and the ASD-LN with high NVIQ group, both for Identical repetition and Target structure, on relative clauses. We thus looked at the types of errors made by the TD4-5 group, the ASD-LN with average NVIQ group, and the ASD-LN with low NVIQ group on relative structures. Results showed that the heterogeneity of errors and the low number of participants involved in each of the three groups made it almost impossible to draw strong conclusions. We report here as an example the errors found in the three groups on object relative clauses (Figure 44). Although no strong pattern could be found among the different errors produced, two considerations can be made. As was found in previous analyses on the ASD-LI and SLI groups, while TD4-5 children tended to make changes to the structure (C substitution, SR for OR), children with ASD-LN engaged in different types of errors (C omission, C addition, Agent/theme reversal, a.o.) that were never found in TD children.

FIGURE 44. ERROR TYPES ON OBJECT RELATIVES (ASD-LN AND TD4-5)

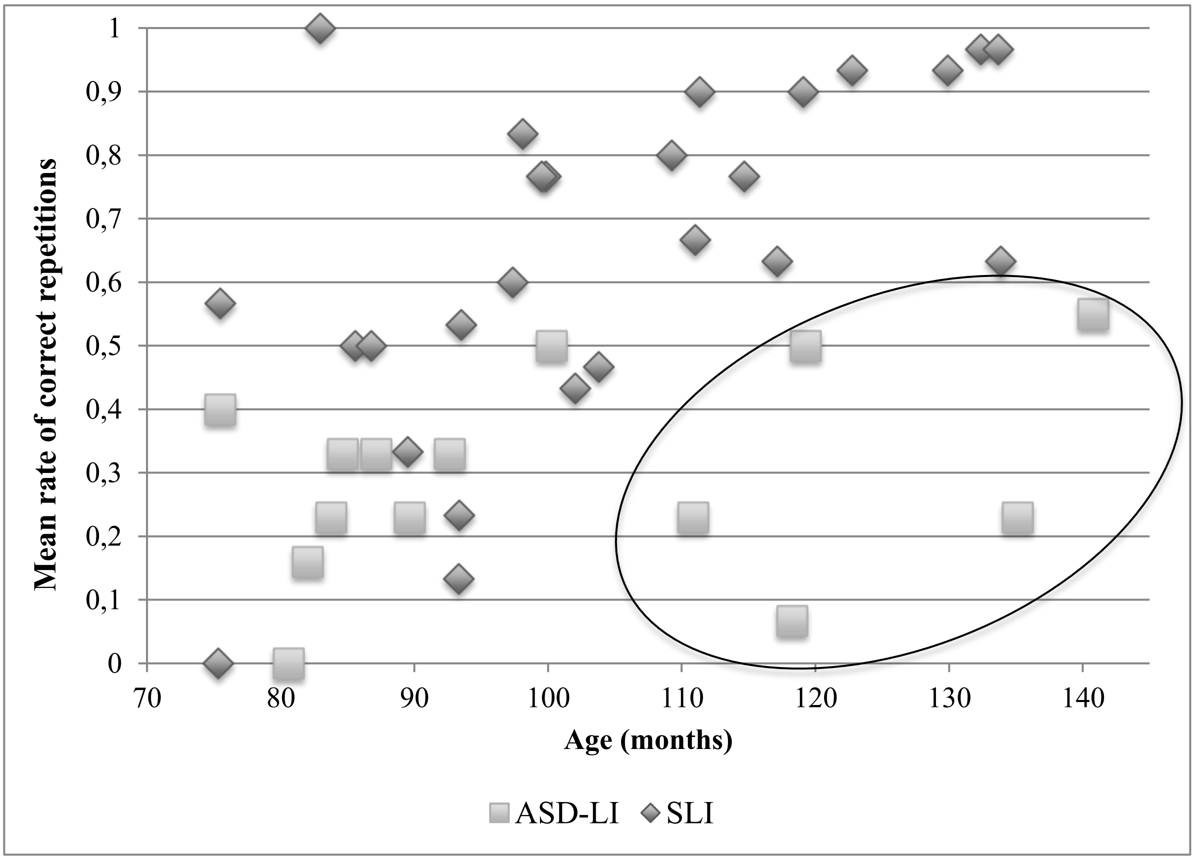


	Subject clitic addition	Tense subtitution	C omission	C substitution	C addition	C misplacement	Coordination	Subject relatives for object relatives	Agent/theme reversal	Embedded object omission	Argument clause for object relative	Fragment + no responses
TD4-5 % errors	0	3	0	7	0	0	0	5	3	0	0	0
ASD-LN average NVIQ % errors	3	3	6	0	6	0	0	0	0	6	3	0
ASD-LN low NVIQ % errors	0	0	0	0	0	8	16	0	0	0	0	8

7.4.4.5 Developmental trajectories

Throughout our quantitative and qualitative analyses of the SR task results, the children with ASD-LI displayed greater impairment than the children with SLI, both in terms of the mean rate of correct repetition and on the number of severely impaired children within the group. Looking back at Figures 24 and 25, we can see that while no child with ASD-LI had over 80% of correct repetition, the children with SLI showed greater heterogeneity, including very high scores (especially on Target structure, with up to 100% correct repetition). We then checked the relationship between age and the mean rate of correct repetition in both language impaired groups and found that while the children with ASD-LI did not show any significant correlation (r_s = .326, p = .255), performing very low regardless of their age, the children with SLI showed a significant improvement (r_s = .594, p = .001) after age 9;0 (Figure 45).

FIGURE 45. INDIVIDUAL PERFORMANCE ON SR OF ASD-LI AND SLI ON AGE. THE BLACK CIRCLE HIGHLIGHTS OLDER CHILDREN WITH ASD-LI WITH SEVERELY IMPAIRED PERFORMANCE ON SR



7.3.4 SR task: conclusions and discussion

The aim of this section was to investigate whether children with ASD-LI would show a phenotypical profile of morphosyntactic abilities similar to children with SLI and whether performance of children with ASD-LN would resemble to performance of younger or age-matched TD children. In order to answer these questions, we looked at the effect of computational complexity on a task of SR specifically built to evaluate linguistic abilities and to disentangle children with language impairment from children showing normal language abilities. Results showed significant computational complexity effects in both children with ASD-LI and children with SLI, as hypothesised in section 7.3.3.3.

However the “phenoptypical” realisation of the children’s impairment presented some striking differences. Sensitivity to a number of syntactic dependencies (e.g. agreement) and/or movement involved in the syntactic derivations of embedded clauses resulted in lower rates of correct repetition on sentences including relative clauses and argument clauses. These results are in line with previous studies investigating complexity effects on embedded structures in French-speaking children with SLI (Jakubovicz & Tuller, 2008; Hamann et al., 2007; Delage et al., 2008; Guasti & Cardinaletti, 2003; Tuller et al., 2006). Moreover, both groups systematically displayed more difficulties on more complex conditions than on less complex conditions for each substructure. Crucially, these differences of performance could not be due to an effect of working memory, since less complex and more complex substructures in each condition did not differ significantly in terms of number of syllables. This means that the lower performance found in more complex conditions was imputable to computational complexity effects only.

Nonetheless, although the general tendency on the task was the same in the two groups, the children with ASD-LI performed significantly lower than the children with SLI on all conditions, suggesting that the children with ASD-LI were more severely impaired than the children with SLI (in line with Taylor et al., 2014). These differences of performance were clearly displayed on all conditions for Identical repetition, while it was reduced to more complex conditions on the Target structure measure. The only exception was represented by *qui* wh-object sentences where both children with ASD-LI and children with SLI performed in line with age-matched TD children. We suppose that this specific condition did not entail a sufficient level of complexity, syntactically speaking (since there was no intervention effect), and also in terms of length (the sentences contained only four words, which might have not been enough to overcome the limit of working memory span), to discriminate the performance of language impaired children from that of TD children.

Was the lower performance of the children with ASD-LI due to some individuals or was it the phenotypical realisation of genuine underlying differences between the two populations? Qualitative error analysis and inspection of developmental trajectories helped us answer this question, indicating that these two groups displayed different behaviours on the task. We found that the children with SLI, when faced with syntactically complex structures, were significantly more likely to make wholesale changes to the syntactic structure than the children with ASD-LI. In line with what was previously found by Novogrodsky & Friedmann (2006) and more importantly by Riches et al., (2010), which was the only study that ran a qualitative error analysis on an experimental task of sentence repetition, the children with SLI tended to transform object relatives into subject relatives, simplifying the structure and reducing computational complexity effects (in this case intervention effects). Crucially, this type of strategy was not limited to relative clauses in our study. Notably, for argument clauses the children with SLI tended to transform finite argument clauses into declaratives (via preposition addition and complementizer omission) and for *quel* wh-object questions they were more prone to produce subject wh-questions instead. Having a variety of structures in our task (including relative clauses) allowed us to both confirm and implement Riches' findings, suggesting that a "structure changing strategy" that aims to overcome computational complexity effects via a simplification of the sentence structure may be in place in children with SLI (in line with Tuller et al., 2012).

The types of errors found in the SLI group were rarely found in the ASD-LI group, indicating that other strategies might have been on the line. Not only was the number of errors higher in the ASD-LI group than in the SLI group, the nature of these errors was also different (contra Riches et al., 2010). Omission errors, which could range from simple determiner omission (an error found in each condition, including monoclausal sentences, in the ASD-LI group) to omission of one or several fundamental elements of the sentences (e.g. omission of the complementizer in embedded clauses, use monoclausals instead of biclausals) were systematically found in our ASD-LI group. The children with ASD-LI did not seem to have the ability to simplify the structure of the sentence, producing a simpler grammatical sentence instead, in contrast to their SLI peers.[10] Further evidence of the fact that different error types were produced by the children with ASD-LI was the Agent/Theme reversal error, which was almost never produced by the children with SLI and the Fragment/No responses strategy which was employed by 64% of the children in ASD-LI group compared to only 20% in the SLI group.

The ASD-LI and SLI groups also differed in terms of the number of children involved in producing errors. In the ASD-LI group, 100% of the children produced at least some errors, while in the SLI group it concerned only 73% children. Crucially, older children with SLI performed much better than their younger peers (although they never reached the same level of TD children). These results suggest that while in our group of children with ASD-LI the syntactic deficit seemed to more severe (in line with Loucas et al., 2008; Modyanova et al., 2017; Sukenink & Friedmann, 2018), children with SLI seemed be less impaired. A possible reason behind the difference between the two studies had to do with the recruitment of the children with SLI. Our children were recruited via the university teaching hospital centre specialized in language and learning disability diagnosis in Tours, where children with SLI receive regular speech language therapy (at least once a week). These results differ from the

[10] We look a posteriori for example at Grammaticality scores (which was not included in the analysis) on embedded clauses in both groups. This analysis showed that while children with SLI displayed a mean rate of 64% of correct grammatical productions on argument clauses (finite and nonfinite) and 72% on relative clauses (subject and object), children with ASD-LI showed respectively 16% and 27% of grammatically correct sentences.

findings in Riches et al. (2010), where the group of adolescents with SLI that was involved showed severely impaired performance on sentence repetition task independently from their age. No direct information about the frequency and the type of language therapy was given by Riches et al. It could be that the children included in their study had received less frequent and less specific language therapy, leading to a more homogenous performance. Another factor that may have contributed to the difference between the two studies could be the heterogeneity of the level of SLI severity. Two children with SLI can share a similar profile (e.g., a weakness primarily in grammatical computation), but differ in the degree to which this primary area is affected (Leonard, 2010). This may yield difference of performance between studies.

Regarding the ASD-LN profiles, there seemed to be a difference between the children with high NVIQ abilities, who tended to perform like age-matched TD6-12 children (in line with our hypothesis and previous findings by Harper-Hill et al., 2013 and Brynskov et al., 2016) and the children with average and low NVIQ who behaved like younger TD4-5 children, showing some selective drop of performance (although not comparable to the one showed by children with SLI and ASD-LI) on more complex structures, such as relative clauses. Again, error analysis showed that while the TD4-5 children were more prone to simplify the structure of the sentence, making wholesale changes (e.g. producing subject relatives instead of object relatives), in line with global findings in the literature (Contemori & Belletti, 2014; Friedmann et al., 2009) the children with ASD-LN engaged in different types of errors (e.g. omission or addition of C). This kind of strategy resembled the behavior of their ASD-LI peers, who produced similar types of errors, albeit much more frequently (both regarding the number of errors and the number of participants involved).

In conclusion consistent with previous research (Conti & Botting Ramsden, 2003, Loucas et al., 2008, Riches et al., 2010, Harper-Hill et al., 2013 Taylor et al., 2014; Brynskov et al., 2016), SR demonstrated some sensitivity as a phenotypic marker of language impairment, in both children with ASD-LI and children with SLI. The difficulties displayed by these children were striking especially given their age, 6-12 years old, suggesting that computational complexity had a strong effect in defining their language impairment. The finding that LITMUS-SR detected a profile of language impairment in the ASD group supports the claim that this task may be a useful clinical marker of language impairment in a variety of different populations with language difficulties. Moreover, the fact that the LITMUS-SR task was built essentially to evaluate the effect of syntactic complexity through a variety of constructions was useful to better describe the difficulties and the errors across computationally complex features in groups of children with language impairment and crucially also in some children with ASD-LN (in line with Tuller et al., 2017; Modyanova et al., 2017). Two subgroups (ASD-LN with low NVIQ and ASD-LN with average NVIQ) in fact did not display language performance in line with age-matched TD children, showing some difficulties on more complex utterances (in particular relative clauses). These results indicate that the SR task could be used as a good endophentoype not only for detecting syntactic impairment in children with ASD-LI, but also for giving a finer-grade picture of syntactic abilities of children with an LN profile, which does not seem to systematically entail completely spared linguistic skills.

One question remains to be addressed: what could be behind the difference of type of errors produced by the children with ASD-LI and the children with SLI? We will discuss the implication of this question in the discussion of the present work (Part IV).

7.4 The NWR task

7.4.1 Participants

The participants were the 37 children with ASD divided on the basis of the four profiles of phonological/NVIQ abilities described in Chapter 6. The children with ASD were compared to the group of 26 children with SLI, and a control group of 42 TD children aged 4-5.

7.4.2 Data analysis

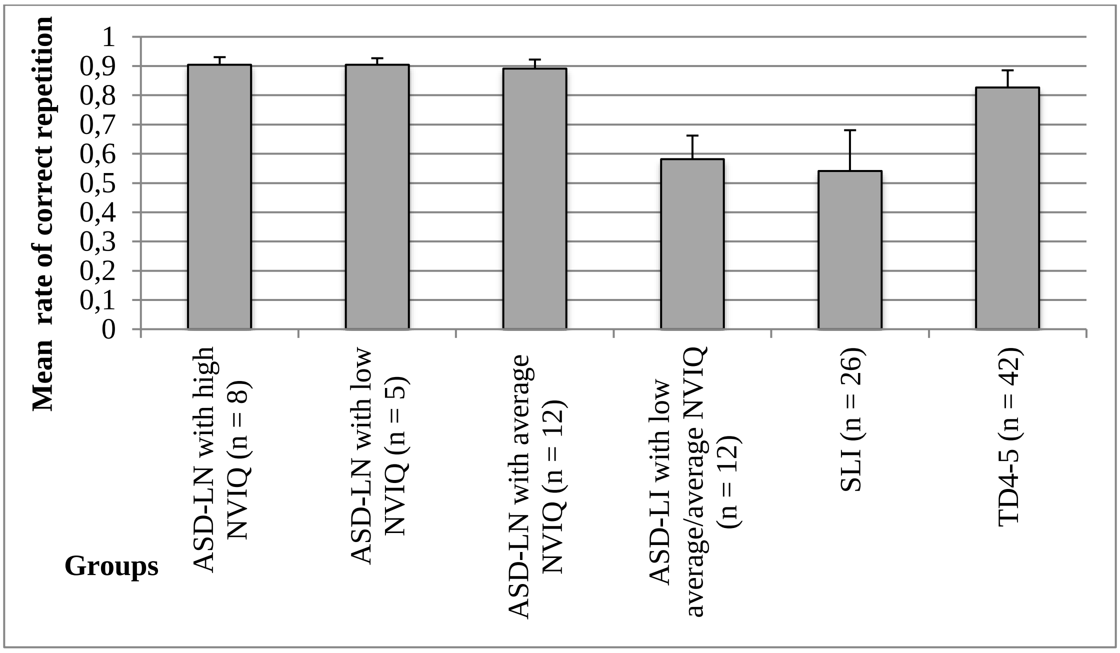
Due to the non-normal distribution of the data (confirmed by the Shapiro-Wilk test) and the small number of children in each subgroup of children with ASD, we followed the same procedure as we did for the SR task (section 7.3.2).

7.4.3 Results on NWR task

7.4.3.1 General performance

We first report the results on the global performance of the five groups on the NWR task (see Figure 46).

FIGURE 46. MEAN RATE OF CORRECT REPETITION ON NWR TASK



The mean rate of correct repetition in the two language impaired groups (ASD-LI and SLI) were similar (and low, around 55%), while the ASD-LN children performed like the TD4-5 group (around 90%). As reported in Table 38, significant differences (after post-hoc correction $p < .003$) were found between the ASD-LI group and all three ASD-LN groups, and between the ASD-LI group and the TD4-5 group.

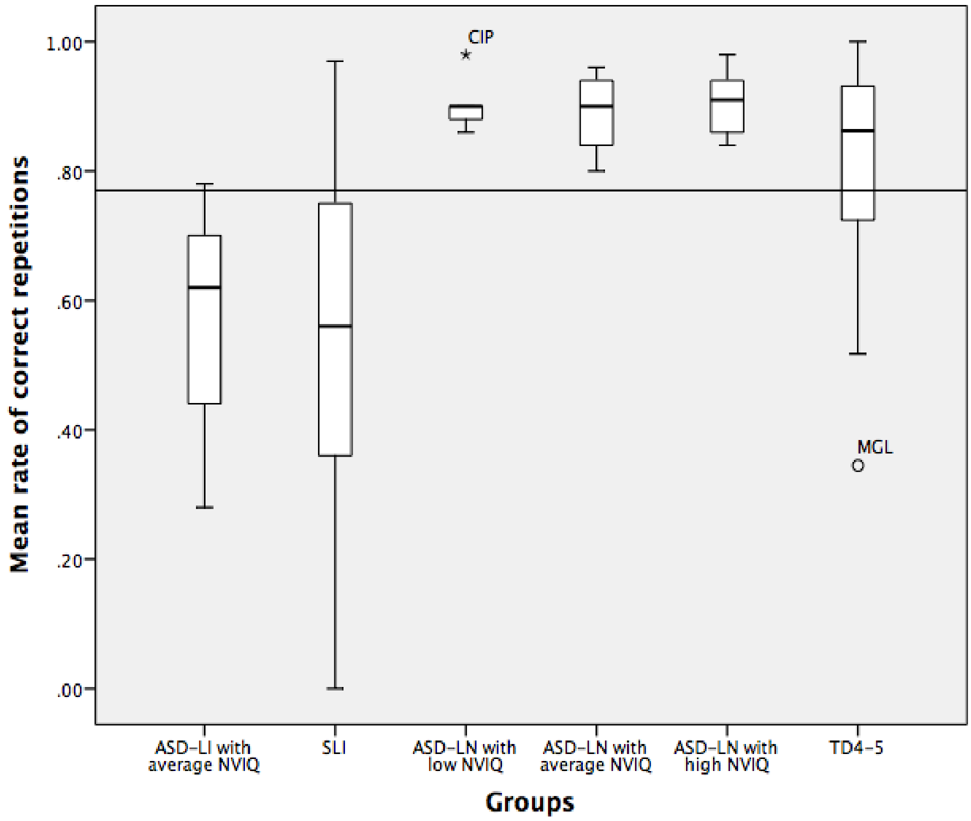
TABLE 38. MANN-WHITNEY INTERGROUP COMPARISONS BETWEEN THE SIX GROUPS FOR NWR

Inter-group comparisons	Correct repetition		
	U	*p*	r
ASD-LN with High NVIQ (*n*= 8) / ASD-LN with low NVIQ (*n*= 5)	13.5	.336	- .26
ASD-LN with High NVIQ (*n*= 8) / ASD-LI with low average - average NVIQ (*n*= 12)	0	**< .001**	- .82
ASD-LN with High NVIQ (*n*= 8) / ASD-LN with average NVIQ (*n*= 12)	44	.754	- .07
ASD-LN with High NVIQ (*n*= 8) / SLI (*n*= 26)	19.5	.001	- .58
ASD-LN with High NVIQ (*n*= 8) / TD4-5 (*n*= 42)	101	.078	- .24
ASD-LN with low NVIQ (*n*= 5) / ASD-LI with low average - average NVIQ (*n*= 12)	0	**.002**	- .76
ASD-LN with low NVIQ (*n*= 5) / ASD-LN with average NVIQ (*n*= 12)	20	.521	- .25
ASD-LN with low NVIQ (*n*= 5) / SLI (*n*= 26)	5	**<.001**	- .57
ASD-LN with low NVIQ (*n*= 5) / TD4-5 (*n*= 42)	49.5	.055	- .27
ASD-LI with low average - average NVIQ (*n*= 12) / ASD-LN with average NVIQ (*n*= 12)	0	**< .001**	- .85
ASD-LI with low average - average NVIQ (*n*= 12) / SLI (*n*= 26)	134	.499	- .10
ASD-LI with low average - average NVIQ (*n*= 12) / TD4-5 (*n*= 42)	44.5	**< .001**	- .58
ASD-LN with average NVIQ (*n*= 12) / SLI (*n*= 26)	31	**< .001**	- .63
ASD-LN with average NVIQ (*n*= 12) / TD 4-5 (*n*= 42)	164	.066	- .24
SLI (*n*= 26) / TD4-5 (*n*= 42)	233	**< .001**	- .47

No difference was found between the ASD-LI group and the SLI group. The three ASD-LN groups did not differ from each other or from the TD4-5 group. All three ASD-LN groups performed significantly better than the SLI group. Finally performance of the SLI group was significantly lower than that of the TD4-5 group.

Figure 47 shows that while almost all children in the ASD-LI and SLI groups performed below the threshold for language impairment (77% of correct repetition, represented by the black line in the graph), all children with ASD-LN performed above this cut-off. In the TD4-5 group some children performed below the threshold. However this was not surprising since the cut-off was established on the basis of a group of older monolingual children (5;6-8;6 years old). Nonetheless, except for "MGL" (who can be considered as an outlier), the TD children did not perform as poorly as the children with ASD-LI and the children with SLI.

FIGURE 47 MEAN RATE OF CORRECT REPETITIONS PRODUCED IN THE GROUPS



7.4.3.2 Individual results

Similarly to what we did for SR, we used the Crawford et al., (2010) t-test for comparison of an individual performance with the control populations. We compared the performance of each child in the ASD-LI group with the performance of the SLI group (Table 39), and the performance of each child in the three ASD-LN profiles with the performance of the TD4-5 control group (Table 40).

TABLE 39. INDIVIDUAL PERFORMANCE OF THE CHILDREN IN THE ASD-LI GROUP COMPARED WITH THE PERFORMANCE OF THE SLI GROUP, USING THE CRAWFORD T-TEST

	Identical repetition	
Child Code	*t*	*p*
FIZ	-1.283	.105
SIM	0.906	.186
NUG	1.547	.067
ASC	1.661	.054
KEV	-0.528	.300
MTH	1.283	.105
LWA	1.547	.067
EVA	0.226	.411
ODI	1.661	.054
EPI	0.377	.354
JOS	0.226	.411
MIR	0.642	.263

TABLE 40. INDIVIDUAL PERFORMANCE OF THE CHILDREN IN THE ASD-LN GROUPS COMPARED WITH THE PERFORMANCE OF THE TD4-5 GROUP, USING THE CRAWFORD T-TEST

		Identical repetition	
Groups	**Child Code**	*t*	*p*
ASD-LN with High NVIQ	MON	0.282	.389
	LEC	0	.500
	YLA	0.282	.389
	TUC	- 0.494	.311
	NOS	0.494	.311
	MUG	0.988	.164
	VOR	0.282	.389
	AVI	1.271	.105
ASD-LN with low NVIQ	ADO	0.494	.311
	ATE	0.777	.220
	LPG	0.777	.220
	YAT	0.282	.389

	CIP	1.271	.105
ASD-LN with average NVIQ	NAF	0	.500
	CUT	0.282	.389
	SCO	-0.494	.311
	ROS	0.988	.164
	EDT	0.777	.220
	EMP	0.282	.389
	GOT	0.494	.311
	ELO	0.494	.311
	GHO	-0.847	.200
	LAT	0.988	.164
	MOI	0.988	.164
	RUG	0.494	.311

No significant difference was found between the individual performance of any child compared to the control groups. This indicates that the children with the ASD-LI profile performed in line with the SLI group and the children in the ASD-LN subgroups performed like the TD4-5 group. In general, then, the observed group performance was confirmed by these individual results.

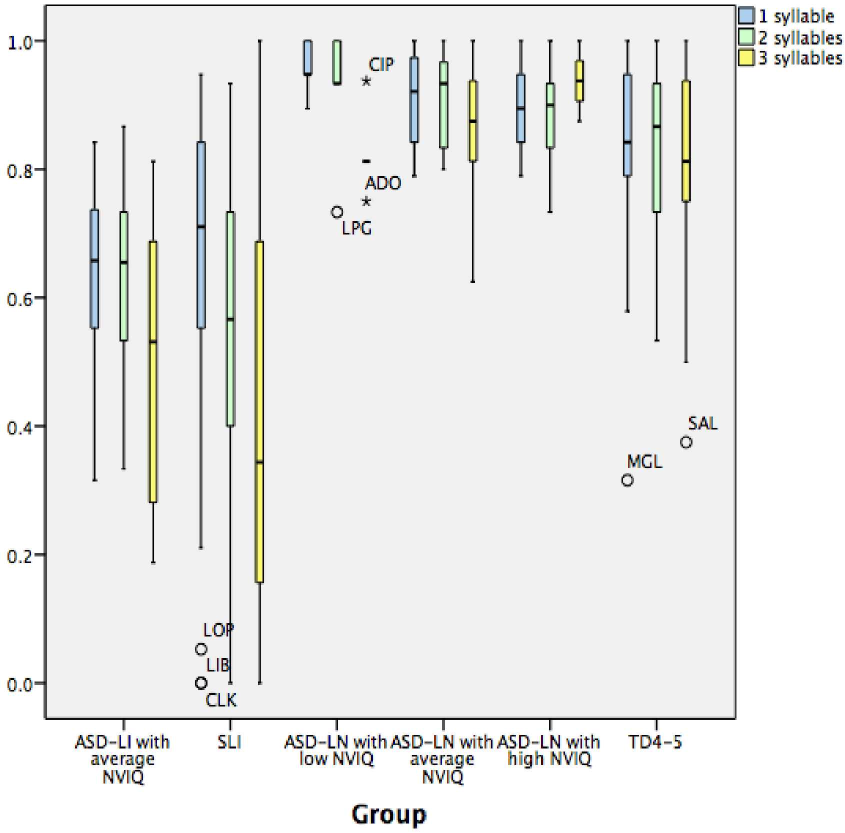
7.4.3.3 Complexity effect on children with language impairment

7.4.3.3.1 Syllable number

The number of syllables is in general considered to have a big impact on repetition tasks. When (non)words have three syllables and more, performance decreases, especially for children with SLI. Our NWR data presented in Figure 48 confirmed the results of these previous studies. All our groups showed a significant decrease in performance between dissyllabic and trisyllabic items, except for the ASD-LN with high NVIQ group who performed better on trisyllabic nonwords (although the difference was not significant). Significant differences between one-, two- and three- syllables items were found only in the two language impaired groups. The children with ASD-LI displayed a difference between one- and three- syllables items (Z = -2.080, p = .038) and between two- and three-syllables items (Z = -2.667, p = .008). No difference was found between one- and two- syllables items in this group. The children with SLI showed significant difference between all three conditions: one- and two- syllables (Z = -2.944, p =

.003), one- and three- syllable items (Z = -3.963, p < .001) and two- and three- syllable items (Z = -2.743, p = .006). The results in the ASD-LI and SLI groups indicate that both groups seemed to be affected by syllable length, with the SLI group sensitive, in addition, to the difference between one-syllable and two-syllable nonwords. Crucially, performance of both groups was never as high as the one displayed by ASD-LN and TD group, not even on one-syllable items.

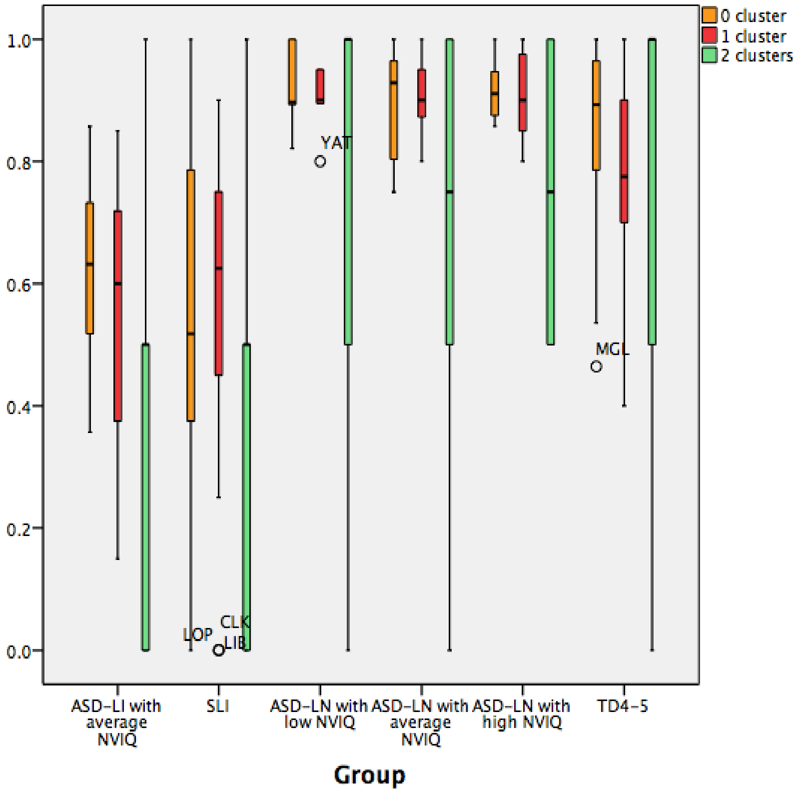
FIGURE 48. MEAN PERCENTAGES OF EXACT REPETITION ON NWR DEPENDING ON THE NUMBER OF SYLLABLES



7.4.3.3.2 Consonant clusters

When two consonant clusters were present within a nonword, there was a large drop in performance for the ASD-LI and SLI children, as shown in Figure 49. In the children with ASD-LI a significant drop of correct repetition was found between items with no consonant cluster, (62.6%) and items with two consonant clusters (37.5%) (Z = 2.492, p = .013). The children with SLI showed a drop in performance both between correct repetition of items with no consonant cluster (38%) and items with two consonant clusters (32.5%) (Z = -1.957, p = .050) and a significant diminution of correct repetitions between items with one consonant cluster (43%) and items with two consonant clusters, (32.5%) (Z = -2.479, p = .013). Performance of the ASD-LN with average NVIQ, ASD-LN with low NVIQ and TD4-5 groups was also impacted by the presence of consonant clusters, but because of high variability in the results within each of these groups, no difference was found between nonwords with one consonant cluster and nonwords with two consonant clusters. For the children with ASD-LI and the children with SLI, the large difference resulting from the presence of a second consonant cluster allowed us to assume that phonological structure was more important than length for detection of impairment, at least in our NWR test, where nonwords were no longer than three syllables. This hypothesis is reinforced by the fact that, again (as found for syllable length), no difference between performance of the children with ASD-LI and of the children with SLI existed regardless of the number of consonant clusters: no consonant cluster ($U(37) = 142$, $p = .442$), one consonant cluster ($U(37) = 152.5$, $p = .646$), two consonant clusters ($U(37) = 164$, $p = .898$).

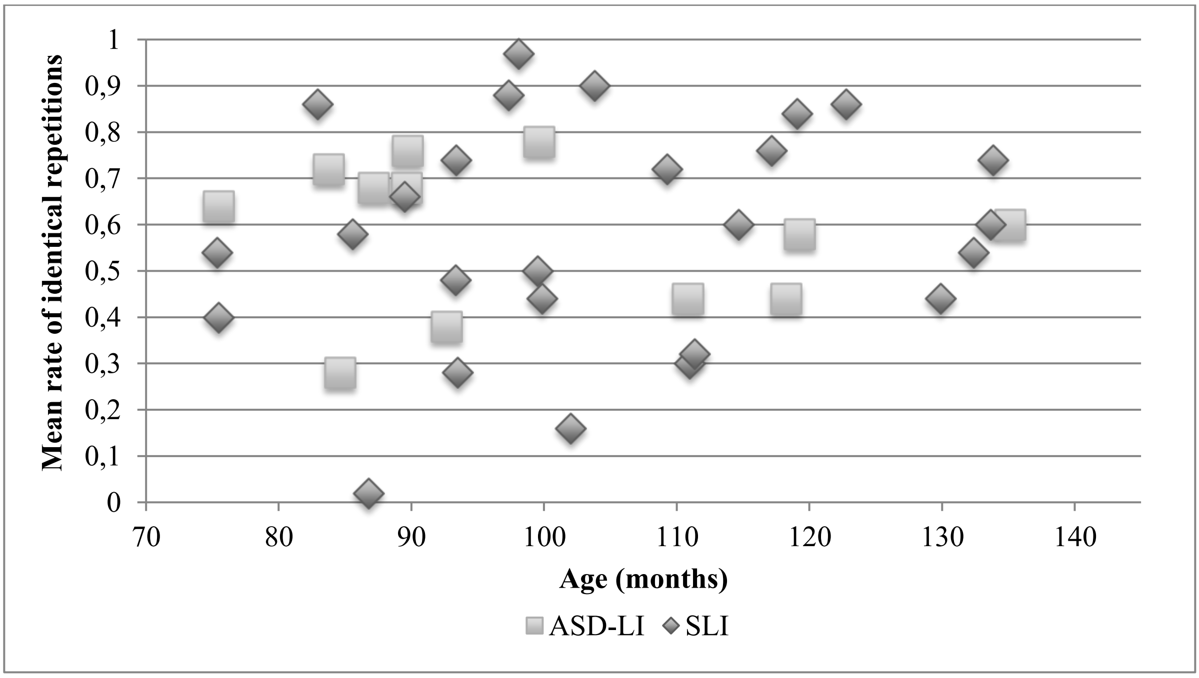
FIGURE 49. MEAN PERCENTAGES OF EXACT REPETITION ON NWR DEPENDING ON THE NUMBER OF CONSONANT CLUSTERS



7.4.3.4 Developmental trajectories

Results on the NWR task suggested that the children with phonological ASD-LI and the children with SLI have a very similar phonological impairment. No difference was found either on general results or on computational complexity measures between the two groups. Figure 50 shows that the two groups performed very similarly, although the children with SLI seemed to display greater heterogeneity of performance (from 0% of correct repetition up to 97%). Similarly to what we did for SR task, we looked at the relationship between age and mean rate of correct performance in both groups (Figure 52). Again, while the ASD-LI group did not display any significant correlation (r_s = -.141, p = 630), there was one in the SLI group (r_s = .527, p = 006). However, Figure 50 shows that the SLI group displayed great heterogeneity of performance, with some younger children performing at ceiling and some older children performing below the cut-off for impaired performance (77%), suggesting that the performance of our group of children with SLI was not systematically related to age, although significant correlation was spotted between these two measures. We will come back to this topic in the next section.

FIGURE 50. INDIVIDUAL PERFORMANCE ON NWR OF ASD-LI AND SLI ON AGE



7.4.4 NWR task: conclusions and discussion

The aim of this section was to investigate whether the children with ASD-LI would display a phenotypical profile of phonological abilities similar to the children with SLI and whether the children with ASD-LN would resemble the TD4-5 children on a task of NWR specifically built to evaluate phonological complexity and controlled for possible effects of working memory.

Comparitive analyses were ran through group results and individual results from a quantitative, qualitative and developmental perspective.

All analyses showed that the children with ASD-LI performed in line with the children with SLI and that the children with ASD-LN performed as the children in the TD4-5 group. The similarity between the children with ASD-LI and the children with SLI has rarely been found in the literature, except for Kjelgaard & Tager-Flusberg, (2001) and Tager-Flusberg, (2015). Crucially, in our study, the ASD-LI and the SLI groups resulted to be similarly affected not only in term of syllable length but also in term of phonological complexity. Both groups displayed a significant drop of correct repetition in three-syllable items compared to two-syllable items and they were both very sensitive to the presence of two consonant clusters in the nonwords, on which they performed very poorly. The children with ASD-LN (except for the children in the ASD-LN with high NVIQ group) and the TD4-5 children showed a slightly effect of syllable length (lower performance on three-syllable items), while they did not display any effect of phonological complexity. These results confirmed the hypothesis that focusing on syllable length may obscure the detection of “real” language impairment, leading to consider as language impaired children that in fact have normal phonological abilities. The use of a NWR task created specifically to test phonological complexity may allow for detection of real phonological impairment in different populations. If we compare these results with the ones on the SR task, where children with ASD-LI seemed to be more severely impaired than children with SLI, the question that remains to be addressed is: to what extent can we talk about similar or different profiles of abilities between ASD-LI and SLI? We will speculate of possible answers to this question in the discussion of the present work (Part IV).

Part III

Bilingual children with ASD

Chapter VIII
Bilingualism and Autism Spectrum Disorder

8.1 Introduction

In Part II of the present work we demonstrated that specific profiles of structural language/NV abilities emerged in monolingual children with ASD when domain specific measures (structural language measures, SR and NWR; NVIQ measures, RPM, Matrix Reasoning and Block Design) were applied. Concentrating on language abilities in these profiles, we showed that monolingual children who displayed impairment on morphosyntactic abilities, the ASD-LI group, did not show the same phenotypical profile of language abilities as children with SLI, performing significantly lower and showing different error typology. We suggested that, at least for morphosyntactic abilities, computational complexity seems to affect monolingual children with ASD-LI differently from children with SLI. For phonology, on the other hand, children with ASD-LI displayed the same phenotypical profile as children with SLI, both on quantitative analysis of their performance (correct repetitions on NWR) and a qualitative analysis of computational complexity effects (syllable length and consonant cluster complexity).

The purpose of Part III of the present work was to investigate whether bilingual children with ASD (BI-ASD) would display the same profiles of structural language/NV abilities as monolingual children with ASD (MO-ASD). In order to answer this question we will apply the methodological procedure used for monolingual children with ASD to a group of bilingual children with ASD and we will investigate whether bilingualism impacts autism, creating separated profiles of structural language/NV abilities from the ones found in monolingual children with ASD. This chapter serves, in addition, as a pilot study to verify whether bilingual children with ASD can be assessed via the same tools (for both language and NV abilities) and the same statistical analyses as monolingual children with ASD. To the best of our knowledge this is the first study to investigate both structural language and NV abilities in bilingual children with ASD.

8.2 Bilingualism and ASD

Bilingualism is an exponentially increasing phenomenon in today's world due especially to increased population flux and globalization. The term 'bilingual' can be applied to any child learning two (or more) languages: a first language (L1), typically spoken at home and in some other contexts, and a second language (L2), mainly learned and used in school (Paradis, 2016). The term 'bilingual' refers to a heterogeneous population in terms of the relative levels of proficiency speakers have in their two languages (Thordardottir et al., 2006). Age of onset (AoO), length of exposure (LoE), quantity and quality of input (rich or reduced input), language status (high or low prestige), L1–L2 language typology, and the sociolinguistic context of language learning can also differ from person to person, which in turn can impact language development and outcomes in each individual.[11]

Considering that the current ratio of autism prevalence in central and southern European countries is roughly 0.1–0.6% (Morales-Hidalgo et al., 2018) - the prevalence in France is estimated to be at 0.36% (van Bakel et al., 2015) - and that the mean rate of bilingual children in Europe has constantly grown during the last ten years from 11 to 20% of the total population (European Commission, 2012), it follows that that the mean rate of children with autism growing up in a bilingual context is also increasing. One example is the retrospective study by Kohl et al., (2008) which reported that out of 47 patients aged 2- to 7- years treated for severe language impairment at Necker Hospital in Paris 16 (34%) came from a bilingual environment. Among these children, 5 (31.3%) were diagnosed to be on the autism spectrum. Although we cannot generalize the results of this study, we can nonetheless conclude that bilingualism is nowadays a reality for many (and an increasing numbers of) children with autism.

For families of children with ASD, questions arise about which language to use with their children. Since difficulties with language and communication are both involved in ASD, the use of two languages seems counterintuitive to many families and raises further questions about how bilingualism could affect these children. Clinicians and SLP are often asked whether children with ASD are capable of learning two languages and becoming bilingual, or whether it would be better for the child's overall (language) development to only hear and speak one language, preferably the societal L2 (Park, 2014). In general, as Hambly & Fombonne, (2012) suggested

[11] In the present work we will not enter the theoretical framework concerning all the different theories about which is the best cutoff of AoO to differentiate children who learn two languages simultaneously from those with later exposure (sequential bilinguals) nor which is the best amount of exposure to a language in order to obtain a L1-like proficiency in L2. For a complete argument on this topic we refer to (Bialystok, 2001)

“there is an urgent need to understand the impact of bilingual exposure on the language development of children with autism since there is a prevalent belief in clinical settings that bilingually-exposed children with ASD may experience additional delays in linguistic development”. Where does this belief come from? In one of the first studies on bilingual children with autism, Baron-Cohen & Staunton, (1994) looked at a group of 10 autistic children (aged 4-14) whose mother’s accent was rated as non-native in English and compared their accents in English with 10 monolingual age-matched children with ASD and 10 bilingual age-matched TD children. Generally, neurotypical children whose parents' native language is other than that of the dominant culture develop an accent that is closer to their peers than to their parents (Cheung & Kemper, 1993). The study examined whether bilingual autistic children would also develop an accent that was closer to that of their peers. Results showed that the participants tended to adopt the speech patterns of their mothers in English more often than their monolingual peers. Their typically developing siblings, in contrast, displayed accents similar to those of their peer monolingual group. Baron-Cohen and Staunton concluded that mothers’ accent (which in the study was generalized as *mother’s language use*) could be considered as a sign of both atypical language acquisition and development in bilingual children with ASD.

However, the ability in displaying L1-like accent cannot be compared to the mastery of abstract and structural rules of a language (linguistic competence in a Chomskyan sense Chomsky, 1965). After Baron-Cohen’s study the literature on bilingual children with ASD came to a halt until the last decade, which has led to a lack of knowledge on the topic which concurred with professionals’ advise to bilingual parents to speak in only L2 (societal language) with their autistic children (Kremer-Sadlik, 2005; but also Paradis, 2016 for an overview on the topic). These kinds of beliefs have been fostered by general beliefs about bilingual development in children with language impairment (see Paradis, 2005 for an overview). It has generally been shown that bilingual TD children display difficulties and error patterns in their L2 which may persist for periods of time that vary from child to child and that tend to overlap with error patterns identified as endophenotypic markers for SLI (see Gathercole, 2002 and Paradis, 2010 for an overview). The majority of studies on French bilinguals have so far focused on structural language abilities as a key direction of research: for morphosyntax, studies have reported striking similarities between bilingual TD children and monolingual French-speaking children with SLI on tense marking (Paradis & Crago, 2000), object clitics (Paradis, 2005) wh-movement (Prévost et al., 2014), and for phonology, on complex syllabic structure (Ferré et al., 2012). Despite a large consensus on endophenotypical resemblances between L2 acquisition and SLI, some qualitative differences have also been reported in the literature, e.g. use of tense morphology

(Paradis & Crago, 2000), strong pronouns (Paradis, 2005), and clausal embedding (Scheidnes & Tuller, 2013).

The general comparability in phenotypical linguistic realisations between bilingual children and children with SLI have generated a new field of research, the study of bilingual children with Specific Language Impairment (BI-SLI), which aims at disentangling the effects of bilingualism from those of SLI, making use of both models of bilingualism and models of language impairment. In recent years, more and more studies have demonstrated that bilingual exposure does not cause additional language delays for children with SLI (Altman et al., 2016; Armon-Lotem & Meir, 2016; Grimm & Schulz, 2014; Marinis et al., 2017; Paradis, 2007; Tsimpli et al., 2016; Tuller et al., 2015). Similar findings have also been reported for bilingual children with Down Syndrome (Feltmate & Rehab, 2008).

In this vein, a few studies have very recently started calling into question the findings by Baron-Cohen and colleagues by exploring the development of language abilities in bilingual children with ASD. As far as we know seven published studies have looked at language abilities in bilingual children with ASD. Six studies investigated vocabulary skills, and one looked at structural language outcomes in this population. The main characteristics of all these studies are reported in Table 41.

In a first study by Hambly and Fombonne (2012) 75 children with ASD aged 36 to 78 months (mean age = 46 months) were assessed on their vocabulary skills via the MacArthur CDI parent report questionnaire in Canada. Children were divided into three groups: monolingual children with ASD (*n* = 30), simultaneous bilingual children with ASD (*n* = 24) and sequential bilingual children with ASD (*n* = 21). The L2 was English or French for all children. The difference between simultaneous and sequential bilingualism was made on the basis of children's first exposure to the L2 (AoO): children were considered simultaneous bilinguals if they had been exposed to their second language before 12 months of age and sequential bilinguals if they had been exposed after 12 months of age. Results showed that when compared on vocabulary skills, no differences were found between the monolingual children with ASD and the bilingual children with ASD, regardless of the AoO.

Similar findings were reported by Ohashi et al. (2012) who assessed 20 Canadian bilingual children with ASD aged 24 to 48 months (mean age 40 months) and 40 age- and NVIQ-matched monolingual peers with ASD. All bilingual children were simultaneous language learners (AoO < 24 months). The children's L2 was French or English, and they were exposed to a variety of L1s. The researchers analysed the children's age of first word and sentence, and their receptive and expressive vocabulary scores from the Preschool Language Scale, PLS-4

(Zimmerman et al., 2002). No significant differences in outcomes/performance on vocabulary knowledge, age of first word and sentence were found between the monolingual children with ASD and those that had bilingual exposure.

In another study by Petersen et al. (2012) vocabulary production (PLS-4) and comprehension (PPVT) were assessed in 14 English (L2)-Chinese (L1) bilingual and 14 English monolingual children with ASD, aged 43 to 73 months (mean age = 59 months). All bilingual children were simultaneous bilinguals (AoO < 36 months). As in previous studies, the bilingual group of children with ASD did not show any disadvantage on English vocabulary production and comprehension performance as compared to their age-matched monolingual peers. Bilingual children's vocabulary production and comprehension were also assessed in Mandarin and found to be comparable to their proficiency in English. No difference in performance was found between the two languages.

Valicenti-McDermott et al. (2013) examined language skills in 40 Spanish/English bilingual children with ASD and 40 monolingual English-speaking children with ASD aged less than three years. Expressive language was assessed via the Rossetti Infant-Toddler Language Scale (Rossetti, 2006), which includes parental report and clinical observation of number of words, presence of word combinations, cooing, babbling, and vocalization. No information was given regarding AoO for English. Results showed there were no significant differences between bilingual and monolingual children with ASD on any measures.

Kay-Raining Bird et al. (2012) collected information via a parental questionnaire about language use and proficiency in a group of 30 bilingual children with ASD with English or French as an L2, aged 8 years old, and a group of 19 age-matched monolingual English- or French-speaking children with ASD. Among the bilingual group, 13 children were considered as simultaneous bilinguals (first exposure to the L2 before age 3) while 17 were sequential bilinguals (first exposure to the L2 after age 3). Again, no difference was found between the bilingual (independently from age of first exposure) and monolingual children on comprehension and production of single words and word combinations.

Finally in recent years, Gonzalez-Barrero & Nadig (2017) examined the impact of bilingualism on verbal fluency in children with ASD. Fifty-two children divided into four groups of 13 TD monolingual children, 13 TD bilingual children, 13 monolingual children with ASD and 13 French-English bilingual children with ASD aged 5 to 10 years (mean age 8;3). Bilingual participants were classified for language dominance (seven children were French dominant and six were English dominant), though no information was given about how this was calculated. Verbal fluency was assessed via the Word association subtest of the CELF-4, or its French version, depending on the child's dominant language. Results suggested that an interaction between bilingualism and ASD allowed the bilinguals with ASD to produce similar number of correct words in the fluency task as their TD peers, while the monolinguals with ASD exhibited poorer performance. This was the only study that showed that being bilingual may improve lexical abilities of children with ASD along with the fact that bilingualism was not reported to have a detrimental effect on vocabulary abilities of children with ASD.

Concerning structural language abilities, as far as we know, only one study has examined morphosyntactic skills in bilingual children with ASD. Baldimtsi et al. (2016) compared six monolingual Greek-speaking children with HFA and six bilingual children with HFA, with Greek as an L2 and multiple languages as L1s (mean age 9;8), to six TD Greek monolingual age-matched children and six TD Albanian/Greek bilingual age-matched children. All bilingual children were equally exposed to both languages from birth. The children's linguistic abilities were assessed via the Edmonton Narrative Norms Instrument, ENNI (Schneider et al., 2005), a narrative exercise of storytelling. Children were evaluated on (1) *lexical diversity* (number of content word types divided by the total number of content word tokens); (2) *language complexity* (number of coordinate and subordinate sentences divided by the total number of simple and complex sentences); and (3) *syntactic complexity* (number of subordinate sentences divided by the total number of complex sentences). Results showed that the two ASD groups did not differ on any of the three scores used to evaluate structural language abilities. On the other hand, monolingual children with ASD differed from TD monolingual group, displaying poorer language complexity, while they did not show any difference in lexical diversity and syntactic complexity. The authors suggested that "the specific difference between the two monolingual groups in language complexity implies that monolingual children with HFA did not link events using coordination or/and subordination to the same extent as their TD monolingual peers, but rather recounted the story by event lists instantiated by simple sentences". These results were not found for bilingual children with ASD, who did not differed neither from monolingual children with and without autism, nor form bilingual TD children. Bilingualism did, however, compensate for children's performance in the macrostructure of narratives, specifically, in story structure complexity and reference use. The authors concluded for a positive effects of bilingualism, especially with respect to children's performance on story telling macrostructural properties.

In summary, most of the available studies on bilingual exposure in children with autism have focused on language measures such as expressive and receptive vocabulary and word combinations in preschool aged children with ASD and found no delays in their language outcomes compared to children with ASD who were not exposed to a second language. Nonetheless these studies have some limitations: (1) Most of the participants (with the exception of Kay-Raining Bird's, Gonzales-Barreiro's and Baldimtsi's studies) were preschool children with ASD; in order to understand linguistic development in bilingual children with ASD we need to include school-aged children; (2) Different criteria regarding AoO have been used for separating children between simultaneous and sequential bilingual groups leading to difficult inter-study comparisons; (3) Children's language abilities were not systematically evaluated via direct assessment, but rather via parental questionnaires. In order to have a more controlled evaluation of language capacities we need to directly evaluate language abilities of children with ASD; (4) Except for Baldimtsi and colleagues, bilingual children with ASD were evaluated only on vocabulary skills. Although this is important, vocabulary outcomes provide only one indicator within the broader context of linguistic development. This approach circumscribes the knowledge of language capacities of bilingual children with ASD to vocabulary, neglecting the underlying structural language abilities of these children; (5) No study has evaluated phonological abilities in bilingual children with ASD.

In the present study we tried to overcome these limitations; a complete explanation of the methodology that was adopted will be presented in section 8.5.

8.3 Cognitive abilities in bilingual children with ASD

Although the effects of bilingualism on cognitive abilities in TD children and adults have been widely investigated, this topic is still controversial. Generally, after Peal & Lambert's (1962) pioneer study on the supposed effect of bilingualism on intelligence, the focus of the debate has shifted from the belief that bilingualism may hinder cognitive development of an individual, to the possibility that individuals raised in bilingual environments enjoy cognitive advantages. More specifically, acquiring a second language has been reported to confer advantages on some specific nonverbal cognitive abilities. It has been reported that on tasks assessing inhibition, monitoring, set-shifting and working memory bilinguals generally perform better than monolinguals (Bialystok & Martin, 2004; Grundy & Timmer, 2017; Hilchey & Klein, 2011; Jarvis et al., 1995; Morales et al., 2013; Okanda et al., 2010). But what about nonverbal cognitive tasks evaluating fluid reasoning and visuospatial abilities? The few studies that have assessed these skills in both bilingual and monolingual children found that when compared on fluid reasoning tasks, e.g. Raven's Progressive Matrices, TD bilingual children did not show any particular advantages over their monolingual peers (Bialystok & Martin, 2004; Hilchey & Klein, 2011; Jarvis et al., 1995). Concerning visuospatial abilities, studies have always evaluated these skills in relation to working memory, notably via visuospatial working memory tasks, concluding that bilingual children outperformed monolingual peers (Blom et al., 2017; Morales et al., 2013). However, since it has been widely shown that bilingualism entails enhanced working memory abilities, we cannot tell whether the high performance of bilingual children on visuospatial working memory tasks is due to the influence of working memory abilities or to enhanced visuospatial skills. No study has selectively compared visuospatial skills (e.g. via tasks such as Block Design) in bilingual and monolingual children. We can hypothesise that, similarly to what has been found for RPM, no difference would occur between these two populations on such tasks. These kinds of NV tests, in fact, are created to minimize verbal requirements and to be as little cultural-dependent as possible. Moreover they are constructed to evaluate logical reasoning, reducing the importance of other specific abilities which may be needed for performing the task (e.g., fine motor or speech skills, inhibition, working memory, etc.). In this sense being bilingual should not create either positive or negative biases on test performance.

To the best of our knowledge no study has specifically investigated cognitive abilities in bilingual children with ASD. Among the seven studies that investigated language skills in bilingual children with ASD, five of them reported corollary results on cognitive abilities, comparing the performance of the bilingual and monolingual children with ASD on specific psychometric tests of cognitive intelligence. Ohashi et al. (2012) and Petersen et al. (2012), which evaluated the NVIQ of their participants via the Merrill-Palmer-Revised Scale of Development (Roid & Sampers, 2004) and the MSEL respectively, found that bilingual children with ASD performed significantly better than their monolingual peers. In Petersen's study, however, differences in demographic factors (e.g., family income) between the groups may have confounded the results, as suggested by Morton & Carlson, (2017). Neither of the two studies indicated whether children with intellectual disability were included in their population samples. Valicenti-McDermott et al. (2013), Gonzales-Barrero & Nadig (2017) and Baldimtsi et al. (2016), on the other hand, did not find any significant difference between cognitive abilities in bilingual and monolingual children with ASD. However, these three studies included only HF children, who were evaluated via different psychometric scales and indexes (Bayley FSIQ, Leiter NVIQ and WISC FSIQ, PIQ and VIQ). Moreover, as we saw for studies of IQ in monolingual children with ASD, results were pooled together in broad bands of intellectual levels which did not share the same cut-offs across studies. Valicenti-McDermott's and Baldimtsi's studies considered children with borderline IQ (from 70 to 79) as HF children, while Gonzales-

Barreiro's study included children in the HF groups only if they had obtained a standard score higher than 80. In conclusion, no study has investigated NV cognitive abilities (fluid reasoning and visuospatial abilities) in samples of bilingual children with ASD which included children with intellectual disability. One of the purposes of the present study was to see whether children with second language exposure taken form the whole spectrum would perform better than monolingual children with ASD on NV tasks of cognitive abilities.

TABLE 41. SUMMARY OF THE KEY FEATURES FOR EACH OF THE SEVEN STUDIES EVALUATING LANGUAGE ABILITIES IN BILINGUAL CHILDREN WITH ASD

Study	Country	Participants	Type of bilingualism	Languages	Evaluation of L2	Evaluation of L1	Cognitive level
Hambley & Fombonne (2012)	Canada	(*n* = 24) simultaneous bilingual children with ASD; (*n* = 21) sequential bilingual children with ASD, (*n* = 30) monolingual children with ASD, all aged 36 to 78 months	Simultaneous (AoO before 12 months) and sequential (AoO after 12 months)	Multiple L1s and English/French L2	Mac Arthur CDI	Mac Arthur CDI for L1s	
Ohashi et al. (2012)	Canada	(*n* = 20) bilingual children with ASD and (n =40) monolingual children with ASD; both groups were aged 24 to 48 months	Simultaneous (AoO before 24 months)	Multiple L1s and English/French L2	PLS-4 (VocP and VocR)		Merrill-Palmer-Revised Scale of Development: NVIQ score. Bilingual children had significantly higher scores than monolingual children
Petersen et al. (2012)	Canada	(*n* = 14) bilingual children with ASD and (*n* = 14) monolingual children with ASD, all aged 43 to 73 months	Simultaneous (AoO before 36 months)	L1 Chinese, L2 English	PLS for VocP and PPVT for VocR	Mac Arthur CDI for Chinese	NVIQ score of MSEL. Bilingual children had significantly higher scores than monolingual children
Valicenti-McDermott et al. (2012)	USA	(*n* = 40) bilingual children with ASD and (*n* = 40) monolingual children with ASD, all aged less than 36 months		L1 Spanish, L2 English	Rossetti Infant-Toddler Language Scale		Bayley Scales FSIQ ≥ 70 standard scores. No significant difference between the groups
Kay-Raining Bird et al. (2012)	Canada	(*n* = 30) bilingual children with ASD and (n = 19) monolingual children with ASD, all aged 8 years	Simultaneous (AoO before 36 months) and sequential (AoO after 36 months)	Multiple L1s and English/French L2	Parental questionnaire		

Gonzales-Barrero & Nadig (2017)	Canada	(*n* = 13) bilingual children with ASD; (*n* = 13) monolingual children with ASD; (*n* = 13) bilingual TD children; (*n* = 13) TD monolingual children, all aged 4 to 10 years	Dominant in French or English	L1 French or English, L2 French or English,	Word association subtest of the CELF-4		Leiter Scale NVIQ score. All children were HF: standard score > 80. No significant difference between the groups
Baldimtsi et al. (2016)	Greece	(*n* = 6) bilingual children with ASD; (*n* = 6) monolingual children with ASD; (*n* = 6) bilingual TD children; (*n* = 6) TD monolingual children, all aged 9;8 years	Simultaneous (AoO from birth)	Multiple L1s, and L2 Greek	Edmonton Narrative Norms Instrument (ENNI)		WISC FSIQ, PIQ and VIQ scores, > 70 standard score. No significant difference between the groups

8.4 Research question

The literature review presented in the preceding section pointed the lack of studies on structural language abilities (just one study on morphosyntax and none on phonology) and NV abilities in bilingual children with ASD. The purpose of Part III of the present work will be to start filling this gap in the literature. To our knowledge this is the first study to investigate structural language abilities and NV cognitive abilities in a group of bilingual children with ASD.

In this study, our main goal was to examine whether bilingualism would have an effect on structural language abilities and NV abilities of children with ASD. In order to investigate this topic we seeked answers to the following question:

Research question: *Do bilingual children with ASD show different profiles of structural language and NV abilities than monolingual children with ASD?*

Our analysis was structured as follows: first we briefly verified that the same measures we selected for evaluating structural language (SR and NWR) and NV abilities (RPM, Block Design and Matrix Reasoning) in monolingual children with ASD were most likely to appropriatley assess structural language and cognitive abilities in our bilingual group as well, and that the bilingual children with ASD did not differ from the monolingual children with ASD. The results of this analysis will be presented in the following sections after a brief description of the participants' characteristics and of the research protocol.

Applying the measures obtained from this analysis, we then investigated which profiles of structural language and NV abilities were present in the bilingual children with ASD and we compared the outcomes with profiles previously detected in the monolingual children with ASD studied in Part II. Finally we checked for the influence of bilingualism on linguistic performance in our BI-ASD group by comparing the scores obtained on both repetition tasks (SR and NWR) with those displayed by the monolingual children with ASD-LI.

8.5 Methodology

8.5.1 Inclusionary criteria and recruitment procedures

Bilingual children were recruited following the exact same criteria we previously detailed in section 4.2 for monolingual children (see (1) - (5)). The only difference with respect to the monolingual participants was that the children had had to be exposed from birth to at least one other language than French. All children meeting these criteria and present at the Tours hospital centre (in the day-care and in the CRA sections) during the recruitment phase of the study were offered the opportunity to participate. Although 30 bilingual children with ASD could have potentially participated to our study, data were gathered for 14 children (13 children recruited from the day-care section and 1 from the CRA). Sixteen children were excluded from the final clinical group for different reasons: for 7 children it was impossible to administrate the protocol because they were too tired, too distracted or too severely impaired (not enough language to perform the entire protocol); 5 parents did not return their consents; 3 children were transferred to other centres and 1 child received a diagnosis other than ASD at the CRA.

8.5.2 Participants characteristics: clinical data

Fourteen verbal bilingual children with ASD, aged 6-12 years old (*M* = 8;9, SD = 20.8, range = 6;5 -12) were recruited from the Autism Centre in Tours. The group was composed of 1 girl and 13 boys. The BI-ASD group with ASD did not differ from the MO-ASD group on age (*U*(50) = 257, *p* = .966). All participants met the criteria for a DSM-5 clinical diagnosis of ASD, confirmed by the ADOS module 1 (few words), 2 or 3, and/or the ADI-R. Clinical information about the participants, including diagnosis and autism severity scores, were collected from the hospital clinical database. Using the ICD-10 criteria retrieved from the database, the group was composed of 8 children with Autistic disorder, 4 children with PDD-NOS and 2 children with Asperger's Syndrome.

Severity of autism scores (the ADOS severity score, the CARS and the ECA-R global score) and developmental factors (age of first word and age of first sentence) were also collected from the clinical database. The BI-ASD group did not differ from the MO-ASD group on measures of autism severity (ADOS: *U*(50) = 129, *p* = .099, CARS: *U*(50) = 138, *p* = .318, and ECA-R: *U*(50) = 93.5, *p* = .288) and on developmental factors (age of fist word: *U*(50) = 159, *p* = .962, and age of first sentence: *U*(50) = 257, *p* = .966).

8.5.3 Research protocol

Children in the BI-ASD group were tested for French via the same research protocol we used for the MO-ASD group (section 4.5). Language abilities were evaluated via standardized language tasks on VocR (ELO or EVIP), VocP (BILO), MorsynR (BILO), MorsynP (BILO) and Phono (BILO), and via the two experimental repetition tasks (SR and NWR). Regarding cognitive abilities, children were assessed via both RPM and a standardized battery of cognitive abilities (when possible, data from the WISC-IV battery were retrieved from clinical records).

All information on bilingualism were collected via a slightly adapted version of a parental questionnaire, the LITMUS-PABIQ, Parents of Bilingual Children Questionnaire (Tuller, 2015) developed during COST Action IS0804 along with the LITMUS-SR and NWR tasks we used in present work (Armon-Lotem et al., 2015). The PABIQ focuses on the child's early language history, the age of contact with the different languages, and the amount and richness of exposure to each language at home and elsewhere (see Appendix 2). These data were used to evaluate the possible relation between bilingualism and structural language/NV outcomes in the BI-ASD group.

8.5.3.1 Language Measures

All 14 children could perform the totality of the language tasks proposed. As explained in section 4.5.1.2, both SR and NWR were specifically designed to target phonological and syntactic structures that have been identified as problematic for children with SLI cross-linguistically. Both of these tools have been demonstrated to be the most discriminatory for identifying language impairment in monolingual children with SLI (Thordardottir et al., 2011 a.o.) and in monolingual children with ASD (Silleresi et al., in press and in Part II of the present work). In bilingual populations, the two repetition tasks have been shown to be good indicators of impairment, significantly distinguishing between SLI and TD (Armon-Lotem & Meir, 2016; de Almeida et al., 2017; Marinis et al., 2017; Tuller et al., 2018; a.o.). Moreover, repetition tasks have been suggested to have a potential advantage in L2 assessment, as they have been demonstrated to be less influenced by LoE in the L2 (Chiat et al., 2013).

Drawing from these findings on monolingual TD children, children with SLI and children with ASD as well as on bilingual TD children and children with SLI, we hypothesise, then, that these tools should also identify language impairment in bilingual children with ASD.

8.5.3.2 Data from the PABIQ

All 14 children were exposed to French and (at least) another language (L1) early on. The bilingual group included five children exposed to Arabic (three to Algerian, one to Moroccan and one to Iraqi varieties), two to Persian, one to Northern Kurdish, one to Vietnamese, one to Russian, one to Turkish, one to Lingala, and two trilingual children, one exposed to English and Dutch and one to English and Persian. In order to identify the children's level of exposure to each language, a Language Dominance Index was calculated from the PABIQ. This index was first based on a language exposure score in each of the child's languages, taking into account AoO (/4), frequency of contact before age four (/4), language early of exposure (/8), LoE (/4), language use at home (/16), language richness which included weekly extra-curricular activities in each language (/14) and language use in school (/5). Adding up the scores on these items (/55) led to the calculation of a language exposure score, for each language. The Language Dominance Index for French was obtained by subtracting the language exposure score in the L1 from the language exposure score for French. Children with a language dominance index between -6 and +6 were classified as "balanced bilinguals", children with a dominance index below -6 "L1 dominant", and children with a dominance index above +6 "French dominant" (Tuller, 2015).

According to this calculation, 10 bilingual children were "French-dominant" and 4 were "balanced bilinguals". The absence of "L1 dominant" participants could have been due to several factors: all children (except for MIM) were born in France; nearly all had been exposed to French from birth (except for MIM and ACH, whose AoO were respectively 4;6 years and 3;0 years); and they all were being schooled and had speech-language therapy (at the day-care hospital in Tours) exclusively in French.

In addition, another score was calculated as a good predictor of language outcomes in bilingual children: Positive Language Development, PLD (/14), which included a score calculated from age of first word, age of first sentence and parental concerns about their child's language in the first years of life.

Following Tuller et al. (2018), some variables (henceforth *bilingual factors*) were selected as the best possible candidates for explaining linguistic performance in bilingual children: AoO, LoE, language early exposure, language use at home, language richness and

PLD. Table 42 presents the main characteristics of each child in the BI-ASD group taken from PABIQ results.

TABLE 42. BI-ASD PARTICIPANTS' CHARACTERISTICS (PABIQ)

Child code	Gender	Age (y;m)	Clinical Diagnosis (ICD-10)	L1(s)	AoO L1 (/4)	AoO French (/4)	LoE L1 (/4)	LoE French (/4)	Early exposure L1 (/8)	Early exposure French (/8)	Lang. use at home L1 (/16)	Lang. use at home French (/16)	Lang. richness L1 (/14)	Lang. richness French (/14)	Lang. Dominance	PLD (/14)
RIV	M	6;5	PDD-NOS	Northern Kurdish	4	4	3.5	3.5	5	7	9	9	7	14	French	6
MIM	M	6;6	Autism	Arabic (Algeria)	4	2	3.5	1	6	1	8	4	0	14	Balanced	4
MVI	M	6;11	PDD-NOS	English / Dutch	4	4	4	4	4	7	7	11	7	13	French	0
LCU	M	7;6	Autism	Vietnamese	4	4	4	4	4	6	4	12	4	14	French	4
AFP	M	7;9	PDD-NOS	Arabic (Iraq)	4	4	4	4	9	5	10	13	5	10	French	0
DIR	M	8;9	Asperger	Russian	4	4	4	4	2	6	3	16	5	14	French	0
ROB	M	8;10	Autism	Arabic (Morocco)	4	4	4	4	5	7	3	16	1	14	French	0
SBI	M	8;11	PDD-NOS	Persian	4	4	4	4	8	8	10	12	6	9	Balanced	2
ACH	M	9;1	Autism	Turkish	4	2.5	4	3.5	8	4	10	7	5	10	Balanced	12
CAT	M	9;9	Autism	Arabic (Algeria)	4	4	4	4	6	7	3	16	0	12	French	8
YVA	F	10;1	Autism	Lingala (Congo)	4	4	4	4	9	7	3	16	2	12	French	14
JON	M	10;9	Autism	English / Persian	4	4	4	4	9	5	6	14	2	12	French	10
TIF	M	11;0	Autism	Arabic (Algeria)	4	4	4	4	6	7	3	16	0	14	French	2
WAP	M	12;1	Asperger	Persian	4	4	4	4	8	8	10	12	6	9	Balanced	2

8.5.3.2.1 Considering bilingualism in standardized language tasks scores

One of the biggest methodological stumbling blocks in assessing bilingual children is that generally standardized tests do not provide bilingual norms. Several studies have now demonstrated that the assessment of bilingual children using diagnostic tests that have been standardized on monolingual speakers may underestimate language capacities of bilingual children (Armon-Lotem & Meir, 2016; Klingner & Artiles, 2003; Paradis, 2005; Tuller et al., 2015). It is a well-known fact that L2 children may take longer to perform like their monolingual peers in their L2 (Geneese et al., 2004 a.o.). Moreover, it is recognized that there are striking similarities between the types of errors displayed by bilingual children in their second language and those shown by monolingual children with SLI (Paradis & Crago, 2000, for an overview). One of the biggest challenges is to determine whether language difficulties observed for a child growing up in a bilingual context are the result of insufficient (early) exposure to this language or of genuine language impairment (see Paradis, 2005). At the same time, some bilingual children may not perform at age-level in their L1 either, due to changes in language dominance (Tuller et al., 2015, a.o.).

These findings have illustrated the need for adapted norms for standardized assessment of bilingual children in order to avoid erroneous conclusions. Guidelines proposed by the American Speech–Language–Hearing Association (ASHA, 2004) stated that, ideally, the assessment of bilingual children should be carried in L1 and L2 using bilingual norms. However, only very few tests provide bilingual norms alongside monolingual norms (e.g. Gathercole et al., 2013). Concerning the tests we used in our research protocol, none was standardized for bilingual children. Ideally we should also have tested the bilingual children's L1(s) on the same language domains and modalities as we did for French. This kind of assessment was possible for only five bilingual children, while the rest could not be evaluated on their L1(s) due to lack of existing standardised tests.[12] This testing impracticability was illustrative of the situation that prevails for most research protocols in the literature (Garthercole et al., 2013).

One way to bypass this obstacle was suggested by COST Action IS0804 (Thordardottir et al., 2015), which recommended an adaptation of the cut-off criteria for

[12] MVI was administrated the CELF battery, while 4/5 Arabic-speaking children were assessed via the ELO-L, Evaluation du Langage Oral chez l'enfant Libanais (Zebib et al., 2017), which is the only standardized test for language assessment in Arabic, normed in Lebanon, but adapted and recorded by native speakers for Algerian, Libyan, Moroccan, and Tunisian varieties. All five children performed very poorly on their L1s both in production and comprehension.

identification of language impairment when monolingually normed tests are applied to bilingual children. Instead of applying the -1.25 SD cut-off (Tomblin, 2011), the following criteria were proposed: -1.5 SD when a bilingual child is tested in the dominant language, -1.75 SD if the child has a balanced bilingualism, and -2.25 when tested in the weaker language. Following these recommendations each child of our BI-ASD group obtained a cut-off for language impairment in both languages (Table 43).

TABLE 43. CUT-OFF FOR LANGUAGE IMPAIRMENT ON BOTH LANGUAGES IN THE BI-ASD GROUP

Child code	Lang. Dominance	Cut-off for L1	Cut-off for French
RIV	French	-2.25	-1.5
MIM	Balanced	-1.75	-1.75
MVI	French	-2.25	-1.5
LCU	French	-2.25	-1.5
AFP	French	-1.75	-1.75
DIR	French	-2.25	-1.5
ROB	French	-2.25	-1.5
SBI	Balanced	-1.75	-1.75
ACH	Balanced	-1.75	-1.75
CAT	French	-2.25	-1.5
YVA	French	-2.25	-1.5
JON	French	-2.25	-1.5
TIF	French	-2.25	-1.5
WAP	Balanced	-1.75	-1.75

We applied the cut-off obtained for French to each of the children's performance on our research protocol, in order to see how many children performed in the impaired range and how many in the normal range. Results will be reported in section 8.6.1 and will be compared with performance on experimental tasks of SR and NWR.

8.5.3.2.2 Considering bilingualism in experimental tasks of SR and NWR

Similar methodological concerns as the ones exposed for standardized testing on the use of monolingual thresholds for identifying language impairment in bilingual children were considered for the experimental tasks of SR and NWR. Recently, Tuller et al. (2018) conducted a Receiver Operating Characteristic (ROC) analysis on SR and NWR the performance of a bilingual sample (*n* = 95) composed of both TD (*n* = 47) and SLI (*n* = 48) L2 French-speaking children, all aged 5- to 9- years (mean age = 6;3). This analysis gave rise to thresholds for language impairment in bilingual populations for both French-LITMUS tasks: 79% of correct repetition for NWR (sensitivity 84% and specificity 77%) and 53% of identical repetition for SR (sensitivity 77% and specificity 76%). Importantly, the bilingual population in Tuller and colleagues' study included 21 L1 dominant children (L1s: Portuguese, Turkish and Arabic), which might have affected the cut-off for language impairment in SR. Since our population did not include L1 dominant children we re-ran the ROC analyses on the bilingual population of Tuller et al. (2018) including only "French-dominant" and "balanced bilinguals" children (*n* = 74). The new **bilingual optimal cut-offs was established at 77% of correct repetition for NWR** (sensitivity 85% and specificity 74%) **and at 75% of identical repetition for SR** (sensitivity 81% and specificity 72%). These new cut-offs adapted for our population sample were used in further analyses.

8.5.3.3 Cognitive measures

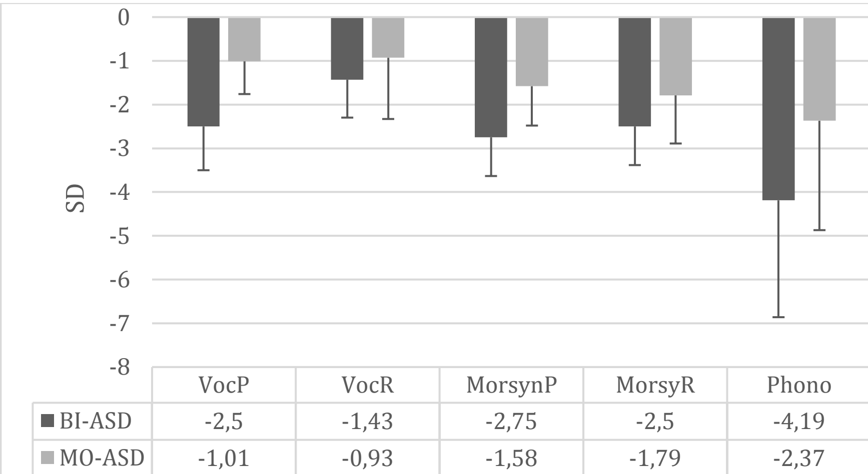
Following the same criteria as in Chapter 4, data for the cognitive level of children with ASD were gathered from the hospital clinical database. Out of the 14 children that were included in our protocol, 10 had been administrated the WISC-IV battery by psychologists. We collected scores for the FSIQ, all four indices, and for the Block Design, Matrix Reasoning and Picture Concepts subtests. The four remaining children were assessed via the EDEI-R. Additional information on the children's NV intelligence was gathered via RPM. The cut-off for intellectual impairment on RPM, WISC-IV (FSIQ, indices and subtests) and the EDEI-R were the same as for monolingual children, < 80 standard scores (section 4.5.1.3). Following the conclusions of section 8.3, we hypothesised that bilingualism would not have any effect on NV cognitive abilities measured on tasks of fluid reasoning and visuospatial abilities.

8.6 Results

8.6.1 BI-ASD performance on standardized language tests

Group results (Figure 51) showed that the BI-ASD group displayed lower performance than the MO-ASD group on all tests of standardized language abilities. However, significant differences were found between the two groups only on measures of VocP ($U(50) = 146$, $p = .022$) and MorsynP ($U(50) = 131$, $p = .047$).

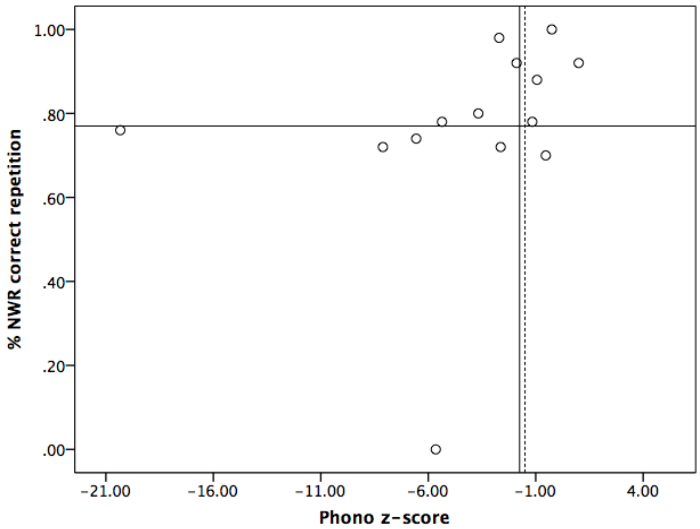
FIGURE 51. Z-SCORES ON STANDARDIZED FRENCH TESTS

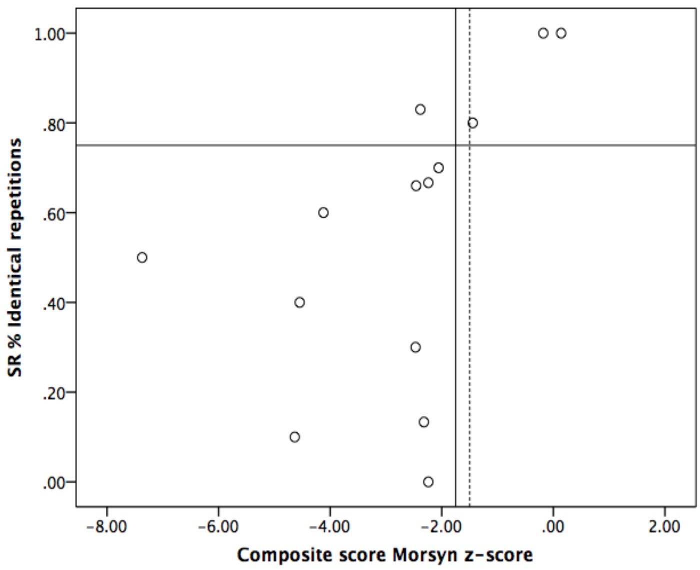


As anticipated in the previous sections (8.5.3.2.1 and 8.5.3.2.2), bilingual children were expected to perform lower than monolingual peers on standardized tests. Following our hypothesis under which experimental tasks of repetition (SR and NWR) were more likely to appropriately assess structural language than standardized tests in our bilingual group, we compared individual performance on repetition tasks and standardized tasks. A composite score of MorsyR and MorsynP was used to represent morphosyntactic abilities on the standardized tests, while for SR the score of identical repetition was adopted. For Phono the z-score from the word repetition task of the BILO was used, while for NWR we adopted the total score of correct repetition. Due to the bilingual status of the participants, we applied the cut-off for language impairment calculated via the Language Dominance Index for each child (Table 43) on standardized tasks and the new cut-off adapted for our bilingual population (section 8.5.3.2.2).

A strong correlation was found between the composite score for morphosyntax and SR ($r_s = .630$, $p = < .016$), while no correlation was found between the phonology score and NWR ($r_s = .487$, $p = .077$). Examination of individual performance revealed some discrepancies, both for morphosyntax and phonology (Figure 52). Two children (in the upper left quadrant of the first graph of Figure 52) performed in the normal range on SR but not on the composite score for morphosyntax, while four children (in the upper left quadrant of the second graph of Figure 52) performed in the normal range on NWR but not on the Phono task. These results mirrored the findings on the MO-ASD group, indicating that for bilingual children with ASD, the SR and NWR tasks seemed to better reveal their structural language abilities as well. No child displayed better performance on standardized tests than on experimental tasks, as showed by the absence of subjects in the lower right quadrant (Fugure 52).

FIGURE 52. COMPARISONS BETWEEN STANDARDIZED LANGUAGE MEASURES NWR AND SR ON THE BI-ASD GROUP





Note: Vertical lines correspond to – 1.75 SD (solid line for balanced bilinguals) and to – 1.5 SD (dotted line for French dominants) cut-off for impairment in standardized measures; horizontal lines correspond to 75% cut-off for impairment in SR and 77% cut-offs for impairment in NWR for bilingual children

To sum up:

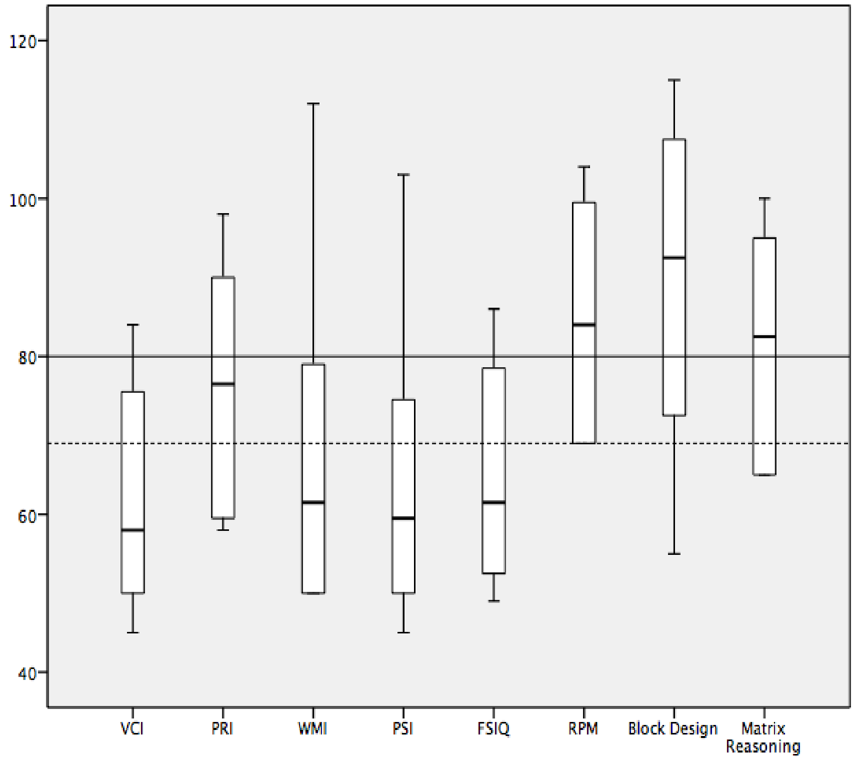
The purpose of the present section was to briefly verify the assumptions that SR and NWR were most likely to be the tasks which would appropriatley assess structural language abilities in bilingual children with ASD, in comparison to standardized tasks evaluating structural language abilities. Our results confirmed this hypothesis.

8.6.2 BI-ASD performance on cognitive tasks

Bilingual children's cognitive level did not differ from that found for monolinguals, either on RPM, or on any score/index of the WISC-IV: FSIQ ($U(50) = 168$, $p = .209$), the four indices of the battery (PRI: $U(42) = 196$, $p = .325$, VCI: $U(42) = 154$, $p = .056$, WMI: $U(30) = 60.5$, $p = .162$, PSI: $U(34) = 83.5$, $p = .166$) and the three subtests of the PRI (Block Design: $U(42) = 130$, $p = .867$, Matrix Reasoning: $U(42) = 93.5$, $p = .164$ and Picture Concepts: $U(32) = 51.5$, $p = .066$).

Similarly to their monolingual peers with ASD, the bilingual children showed better performance on RPM, Block Design and Matrix Reasoning tests than on all other measures of cognitive abilities (Figure 53); significant differences were found between RPM and VCI scores ($Z = -2.482$, $p = .013$), RPM and FSIQ scores ($Z = -2.273$, $p = .023$), and Block Design and VCI scores ($Z = -2.310$, $p = .021$). Moreover, no significant correlations were found between the measures of RPM, Block Design and Matrix Reasoning.

FIGURE 53. BI-ASD GROUP PERFORMANCE ON PSYCHOMETRIC MEASURES OF COGNITIVE ABILITIES



Note: The cut-off for performance below the norm was < 80 standard score (indicated by the solid line) and for mild/moderate intellectual impairment it was < 70 standard score (indicated by the dotted line)

Finally, results showed that no correlation was present between NV cognitive scores on RPM, Block Design and Matrix Reasoning and language abilities measured via SR and NWR, indicating that like for the monolingual children, the NV measures selected were not linked to verbal abilities in the bilingual group. Measures selected were not linked to verbal abilities. All these results are shown in Table 44.

TABLE 44. SPEARMEN'S CORRELATIONS BETWEEN NV MEASURES AND LANGUAGE MEASURES

		RPM	Block Design	Matrix Reasoning	NWR	SR
RPM	r_s		.624	.095	.380	.496
	p		.072	.627	.200	.085
	N		10	10	14	14
Block Design	r_s	.624		.010	.536	.333
	p	.072		.957	.137	.073
	N	10		10	10	10
Matrix Reasoning	r_s	.095	.010		.121	.381
	p	.627	.957		.756	.352
	N	10	10		10	10

8.7 Conclusions

In conclusion, our analysis of BI-ASD group performance on standardized tests of language abilities and on psychometric tasks of cognitive abilities showed that no difference occurred between the bilingual children with ASD and the monolingual children with ASD. Following these findings, we assumed that the methodological conclusions of Chapter 5 would also be valid for the present analyses on the bilingual children with ASD. We chose, then, to use the SR and NWR tasks for the evaluation of structural language abilities and RPM, Block Design and Matrix Reasoning for the evaluation of NV cognitive abilities. These scores were used for the identification of structural language/NV ability profiles in our BI-ASD group (Chapter 9).

Chapter IX
Structural language/NV cognitive ability profiles in bilingual children with ASD

9.1 Introduction

The main aim of our analysis was to investigate whether bilingual children with ASD (BI-ASD) would display profiles of structural language/nonverbal abilities similar to their monolingual peers (MO-ASD). In order to answer this question, we will concentrate our analyses first on the identification of profiles in the bilingual children with ASD in relation to the profiles of the children in the MO-ASD group. We will see whether the bilingual children with ASD displayed the same profiles of language/NV abilities as the ones found in the monolingual children with ASD. We then will compare the results on structural language abilities (both on SR and NWR tasks) to the results from the Mo-ASD children. This analysis aimed to determine whether the bilingual children with ASD-LN had language skills similar to those of monolingual children with ASD-LN and if the bilingual children with ASD-LI displayed language abilities like the monolingual children with ASD-LI.

9.2 Structural language / NV ability profiles in bilingual children with ASD

9.2.1 Methods

9.2.1.1 Participants

The results presented in the first part of the present chapter include only the ten bilingual children with ASD who were assessed through the WISC-IV task (age M = 9.42, SD = 18.4). According to the ICD-10 criteria, the group included 5 children with Autistic disorder, 3 children with PDD-NOS, and 2 children with Asperger’s Syndrome. The group was composed of 10 boys. Analogously to what we have done for the monolingual children with ASD the performance of the four children assessed via the EDEI-R was analysed *a posteriori* by matching their profiles to the ones detected in our analysis of structural language/NV abilities on the other 10 children.

9.2.1.2 Materials and procedure

Following Chapter 8, the following analyses were based exclusively on the scores from the two experimental repetition tasks targeting specific aspects of structural language (SR and NWR), and the scores from RPM and the two subtests of the PRI, Block Design and Matrix Reasoning, which assessed cognitive abilities..

9.2.1.3 Data analysis

Following the results of Chapter 6 which showed that there were different profiles of structural language and NV abilities in monolingual children with ASD, we ran two separate cluster analyses for the identification and the description of bilingual profiles, one for morphosyntactic abilities and one for phonological abilities. Our analyses on the bilingual group were then integrated to our previous results on monolingual children. An explanation of the methodology behind these analyses is given in the next section.

9.2.2 Results

9.2.2.1 Cluster analyses

Taking as our point of departure our research question, which asked whether bilingual children with ASD display profiles of structural language and nonverbal abilities similar to their monolingual peers, we decided to further implement the exploratory cluster analysis of Chapter 6 on monolingual children, adding bilingual children as new data. However, we did not run a new cluster analysis including both ASD groups (MO-ASD and BI-ASD). This would have probably changed the distribution of monolingual children into ability profiles, since with the inclusion of bilingual data, the algorithm (K-means method) would have partitioned subjects from scratch, calculating a new centroid for each cluster. Instead, using an algorithm of prediction (Leisch, 2006), we calculated the probability for each new subject of the BI-ASD group to fall within the already existing clusters created on the basis of the MO-ASD group. This algorithm allowed us to assign the new observations (the participants in the BI-ASD group) to the existing clusters based on their existing centroids (K-means analyses of Chapter 6). If a subject did not fall in any of the existing clusters, the algorithm considered it as an outlier.

We ran two separate analyses, one for morphosyntactic/NV profiles (on the measures of SR, RPM, Block Design and Matrix Reasoning) and the other for phonological/NV profiles (on the measures of NWR, RPM, Block Design and Matrix Reasoning). The result of these analyses returned K models to match K clusters, one model per cluster, predicting whether each new subject was likely to belong to each existing cluster. Here we present only the graphical realisation of the predicting algorithm results. Figure 54 shows the results of the cluster analysis for morphosyntactic/NV ability profiles and Figure 55 for phonological/NV ability profiles. For ease of identification, the bilingual children with ASD were graphically represented in boldface. Outliers appeared in red.

FIGURE 54. BI-ASD DISTRIBUTION ON THE FOUR PROFILES OF MORPHOSYNTACTIC/NV ABILITIES IDENTIFIED IN THE MO-ASD GROUP

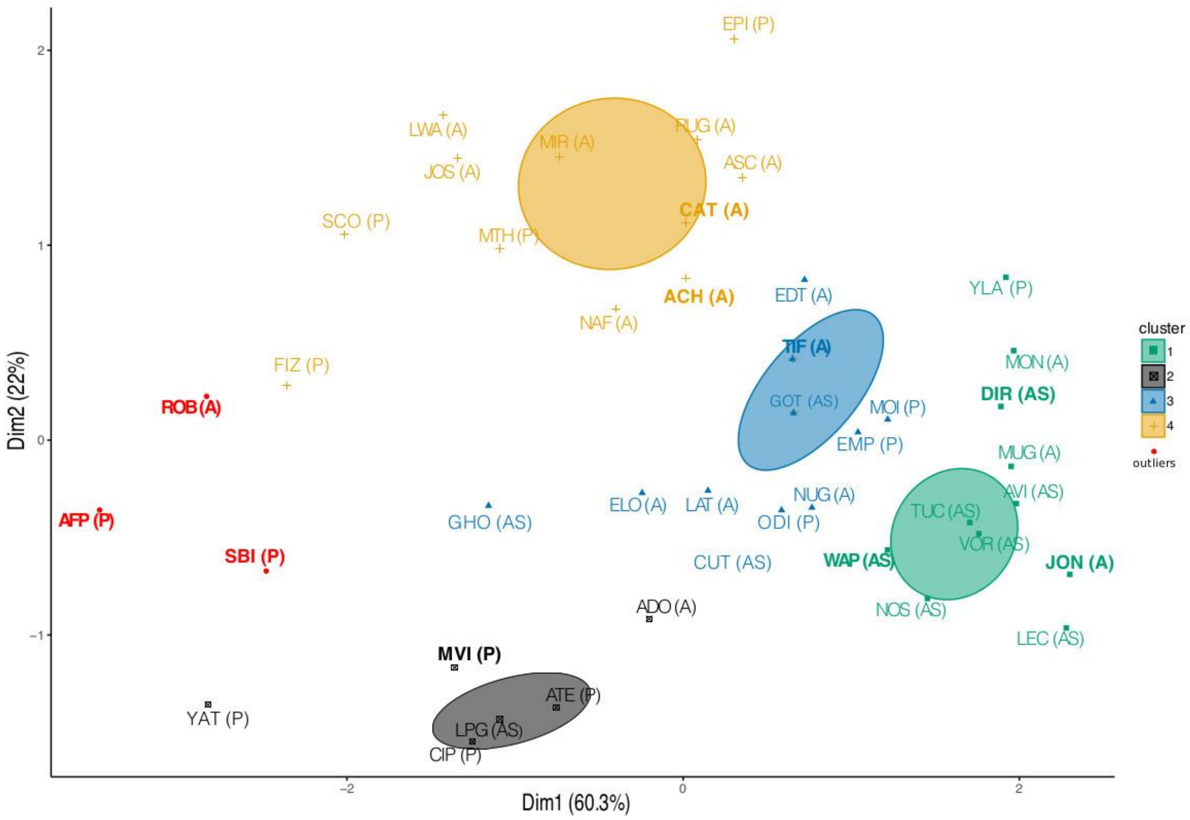
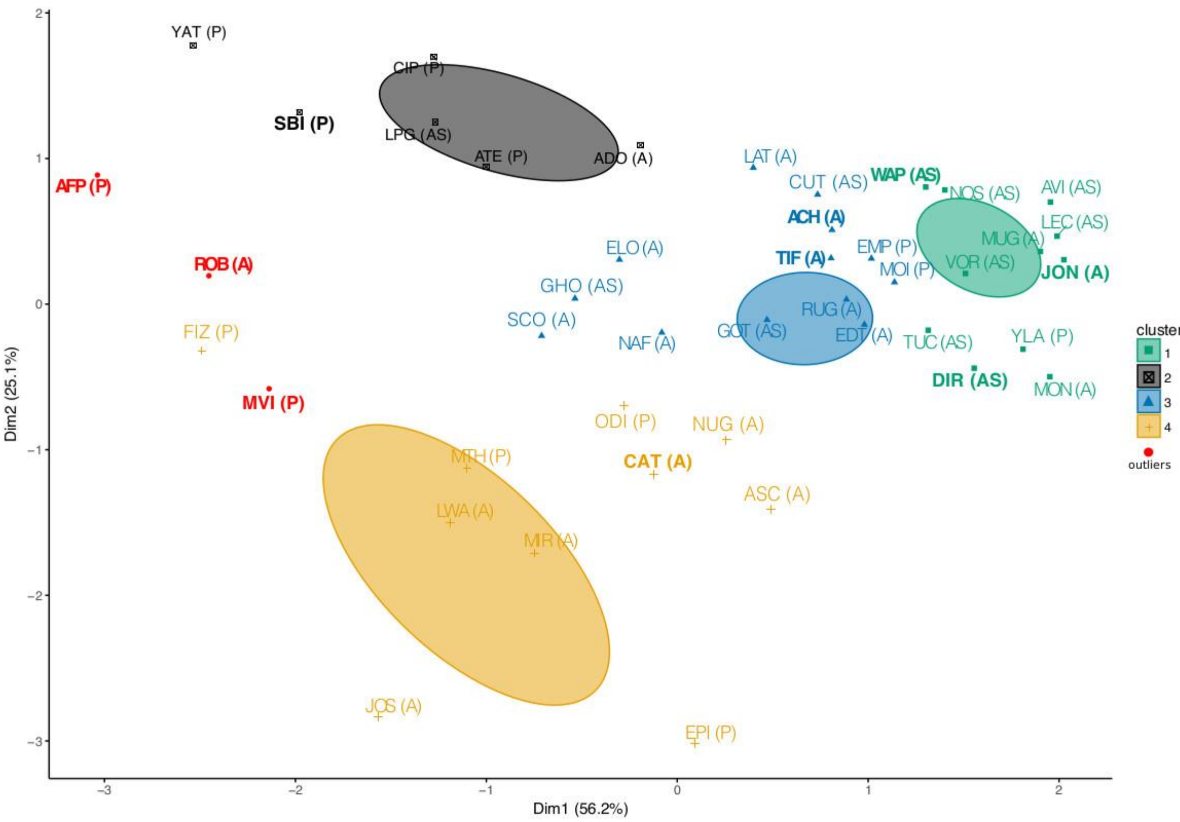


FIGURE 55. BI-ASD DISTRIBUTION ON THE FOUR PROFILES OF PHONOLOGICAL/NV ABILITIES IDENTIFIED IN THE MO-ASD GROUP



The algorithm automatically distributed 6 children into the existing monolingual clusters for both analyses, and considered 4 children as outliers in at least one analysis. Among the six children collocated in the existing clusters, five showed the same profiles for both morphosyntax and phonology: three (WAP, DIR, JON) ended up in the ASD-LN with high NVIQ profile, one (TIF) in the ASD-LN with average NVIQ profile, and one (CAT) in the ASD-LI with average NVIQ profile. The remaining child (ACH) showed a performance similar to the LI with average NVIQ profile, on morphosyntax, and a performance similar to the LN with average NVIQ profile, on phonology. Concerning the four outliers, the algorithm considered the profiles of two children (ROB and AFP) as not comparable with any existing profiles of ability previously detected in monolingual children with ASD, both for morphosyntax and phonology. Regarding the other two children, MVI fell in the ASD-LN with low NVIQ profile concerning morphosyntax, while he was considered an outlier for phonology. Conversely, SBI was considered as belonging to the ASD-LI with low NVIQ profile for the phonological analysis, while he ended up as an outlier in the morphosyntactic analysis.

In general we can conclude that despite the small number of participants, the children of the BI-ASD group were found to fall into all four clusters previously detected in monolingual children with ASD for both morphosyntactic/NV abilities and phonological/NV abilities. In Table 45 we report individual performance of all ten children of the BI-ASD group and their profiles of morphosyntactic/NV and phonological/NV abilities.

TABLE 45. INDIVIDUAL SCORES OF THE TEN BILINGUAL CHILDREN WITH ASD

Child Code	NWR (%)	SR (%)	RPM (standard score)	Block Design (standard score)	Matrix Reasoning (standard score)	Phono/NV profile	Morsyn/NV profile
DIR (AS)	92	100	95	125	110	ASD-LN with high NVIQ	ASD-LN with high NVIQ
JON (A)	88	100	125	100	110	ASD-LN with high NVIQ	ASD-LN with high NVIQ
WAP (AS)	100	80	104	105	100	ASD-LN with high NVIQ	ASD-LN with high NVIQ
TIF (A)	98	83	90	115	95	ASD-LN with average NVIQ	ASD-LN with average NVIQ
CAT (A)	76	40	95	110	90	ASD-LI with average NVIQ	ASD-LI with average NVIQ
ACH (A)	92	66	104	100	95	ASD-LN with average NVIQ	ASD-LI with average NVIQ
MVI (P)	70	80	69	85	75	outlier	ASD-LN with low NVIQ
SBI (P)	80	60	69	75	65	ASD-LN with low NVIQ	outlier
AFP (P)	76	13	69	55	65	outlier	outlier
ROB (A)	72	10	78	70	65	outlier	outlier

Note: Scores below the thresholds are highlighted in grey;
Cut-off for low performance on SR was established at < 75 % of correct repetition (adapted from Tuller et al., 2018);
Cut-off for low performance on NWR was established at < 77 % of correct repetition (adapted from Tuller et al., 2018);
Cut-off for low NVIQ was established at < 80 standard score.

Some children (n = 4) fell outside of the previous classification, displaying profiles of abilities apparently inexistent in the MO-ASD group. Why were these children considered as *outliers*? The "outlier" status of MVI, AFP, ROB and SBI was due to the fact that these children displayed impaired abilities in one or both structural language domains, in addition to having impaired NV abilities. This kind of profile, which we henceforth call "ASD-LI with low NVIQ" was not detected as a profile *per se* in the MO-ASD group. Nonetheless, it was found in two children, FIZ (for both morphosyntax and phonology) and YAT (for morphosyntax). The fact that only 2 out of 33 children displayed such a profile in the MO-ASD group was not sufficient to lead the K-means algorithm to group these children into a separate cluster. However, if we look at the new distribution of both monolingual and bilingual children with ASD-LI and low NVIQ in both cluster analyses (Figures 54 and 55), we can clearly see that ROB, AFP, SBI, FIZ and YAT may constitute a fifth profile in the morphosyntactic/NV analysis, as well as a fifth profile in the phonological/NV analysis.

To sum up, we can conclude that the bilingual children with ASD displayed profiles of structural language/nonverbal abilities similar to their monolingual peers. In addition our results showed the existence of another profile of structural language/NVIQ abilities, the ASD-LI with low NVIQ profile.

9.2.2.2 Children assessed with the EDEI-R psychometric test

Four bilingual children were assessed via the EDEI-R psychometric test and for statistical reasons they were not included in the previous analysis. Similarly to what we did in section 6.4 for monolingual children, we incorporated these four children into the profiles detected via the cluster analyses in section 9.2.2.1. On the basis of individual scores for SR, NWR, RPM and the NV score of EDEI-R (Table 46), we could include RIV in the ASD-LI with average NVIQ profile, LCU in the ASD-LN with average NVIQ profile for phonology and in the ASD-LI with average NVIQ profile for morphosyntax, and YVA and MIM in the ASD-LI with low NVIQ profile for both morphosyntax and phonology, bringing the total number of children in this new fifth profile to eight.

TABLE 46. INDIVIDUAL SCORES OF FOUR CHILDREN ASSESSED VIA THE EDEI-R BATTERY

Child code	NWR (%)	SR (%)	RPM (standard score)	NV score (standard score)	Phono/NV profile	Morsyn/NV profile
RIV (P)	72	30	103	103	ASD-LI with average NVIQ	ASD-LI with average NVIQ
MIM (A)	0	0	78	79	ASD-LI with low NVIQ	ASD-LI with low NVIQ
LCU (A)	78	70	110	104	ASD-LN with average NVIQ	ASD-LI with average NVIQ
YVA (A)	74	50	69	30	ASD-LI with low NVIQ	ASD-LI with low NVIQ

9.2.2.3 Conclusions

In conclusion, this section has shown that bilingual children with ASD displayed the same profiles of structural language and NV abilities as monolingual children with ASD. After separating performance on morphosyntax and phonology, bilingual children distributed homogeneously into the four profiles of structural language and NVIQ abilities in both analyses. Differently from monolingual results, a clear ASD-LI with low NVIQ profile, displayed by six BI-ASD children, emerged in the bilingual sample. The existence of this profile, which we hypothesised also for monolingual children on the basis of previous findings in the literature (Joseph et al., 2002; Kjelgaard & Tager-Flusberg, 2001; Tuller et al., 2017), emerged once a minimum number of children with impaired NVIQ was included in the population sample. We believe that the higher number of ASD-LI with low NVIQ profiles among the bilingual children was not attributable to their bilingual status but to recruitment chance. Evidence for this hypothesis was the fact that the two monolingual children with ASD-LI and low NVIQ that were found in the MO-ASD group clearly showed a similar profile of language/NV abilities as these bilingual children. In addition, similarly to the monolingual children with ASD, low NVIQ did not systematically entail impaired language performance in the BI-ASD group. Two profiles of normal language abilities were found: one including children with normal NVIQ and one with low NVIQ. The existence of a double dissociation like the one found in the ASD-LI with average NVIQ and the ASD-LN with low NVIQ profile indicates that even bilingual children with ASD can display impaired language abilities in presence of spared nonverbal intelligence, as in SLI, and spared language abilities in the presence of impaired nonverbal intelligence, which corresponds to the profile found in the rare condition of the polyglot Savant *Christopher* (who could read, write and communicate in any of 15 to 20 languages) (Smith & Tsimpli 1995).

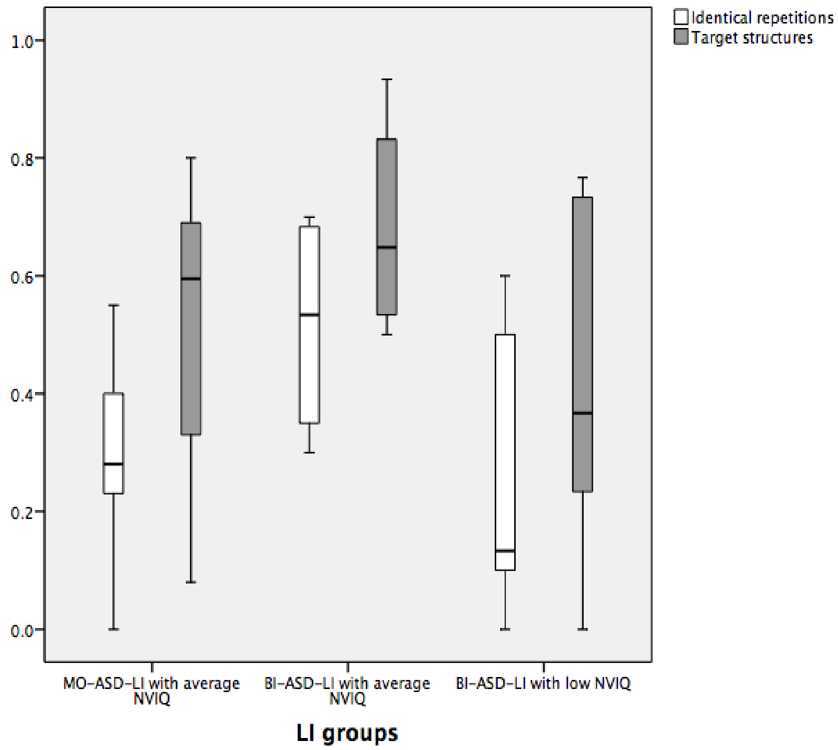
The next step in our analysis will be to investigate language impairment in both BI-ASD and MO-ASD groups. We will concentrate on performance in SR and NWR in order to compare bilingual children with ASD-LI (BI-ASD-LI) with monolingual children with ASD-LI (MO-ASD-LI). The objective of this analysis will be to see whether the phenotypical realisation of language impairment in bilingual children would resemble the one found in monolingual children with ASD. We will not report on children with a BI-ASD-LN profile since no difference was found with the MO-ASD-LN profiles. All children in the three LN profiles did not differed from their monolingual peers neither on SR task (Identical repetition and Target structure scores, substructures and error types) nor on NWR (syllable length and consonant clusters). Results of the BI-ASD-LN profiles are reported in Appendix 3.

9.3 Phenotypical LI profile in bilingual children with ASD

In this section we will briefly analyse the performance of bilingual children with ASD who displayed impaired performance on SR and/or NWR tasks and we will compare their scores with the ones showed by monolingual children with ASD-LI.

Nine children showed impaired performance on the SR task: four (ACH, CAT, LCU, RIV) were included in the ASD-LI with average NVIQ profile, while five (ROB, AFP, SBI, MIM, YVA) were included in the ASD-LI with low NVIQ profile. Figure 56 shows general group performance for Identical repetition and Target structure.

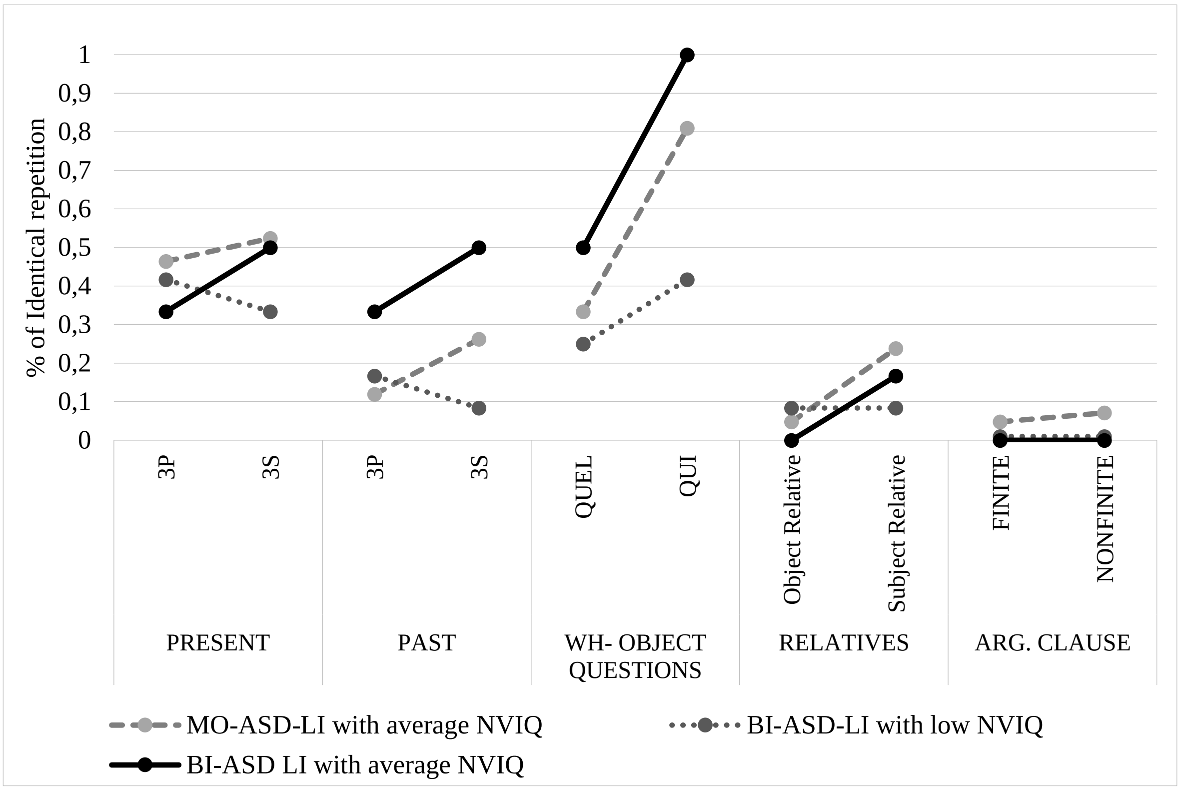
FIGURE 56. MEAN RATE OF CORRECT IDENTICAL REPETITION AND TARGET STRUCTURES IN LI GROUPS



No significant difference was found between the MO-ASD-LI with average NVIQ group and the BI-ASD-LI with average NVIQ group on Identical repetition ($U(17) = 30$, $p = .674$) and Target structure ($U(17) = 28$, $p = .547$) scores. The BI-ASD-LI with low NVIQ group performed below both the MO-ASD-LI group (Id rep: $U(18) = 32$, $p = .022$, Targ. struct.: $U(18) = 29$, $p = .045$) and the BI-ASD-LI with average NVIQ group (Id rep: $U(8) = 12$, $p = .012$, Targ. struct.: $U(8) = 11$, $p = .018$). In contrast to the two bilingual groups, the MO-ASD-LI group showed a significant difference between Identical repetition and Target structures scores ($Z = -3.296$, $p = .001$).

Comparing performance on the different substructures of the SR task (Figure 59), although no significant differences were found between the three profiles on any measures, we can see that the children in the BI-ASD-LI with low NVIQ group behaved atypically on the task. This was especially true for monoclausal SVO sentences. Moreover, children with ASD-LI and low NVIQ did not show the usual effect of computational complexity which predicts better performance on less complex substructures than on more complex substructures in each condition, contra the other two ASD-LI groups. Similar results were found for Target structure scores.

FIGURE 57. LI GROUPS PERFORMANCE ON SUBSTRUCTURES IN THE SR TASK



Why was performance in the BI-ASD-LI with low NVIQ group different from the other two LI groups? Looking at individual performance, the five children included in this group (SBI, YVA, ROB, AFP, MIM) showed unusual behaviours on the task. Two children displayed a significant difference between the mean rate of correct repetitions in the first half (items 1-15) and second half (items 16-30) of the task: SBI's performance on correct repetition improved from 30% to 75% and YVA's improved from 40% to 80%. We hypothesise that factors other than linguistic abilities, such as inattention, may have played a role in these children's performance. Both were, in fact, capable of correctly repeating both less complex and more complex structures in the second half of the task, while they performed very poorly in the first half.

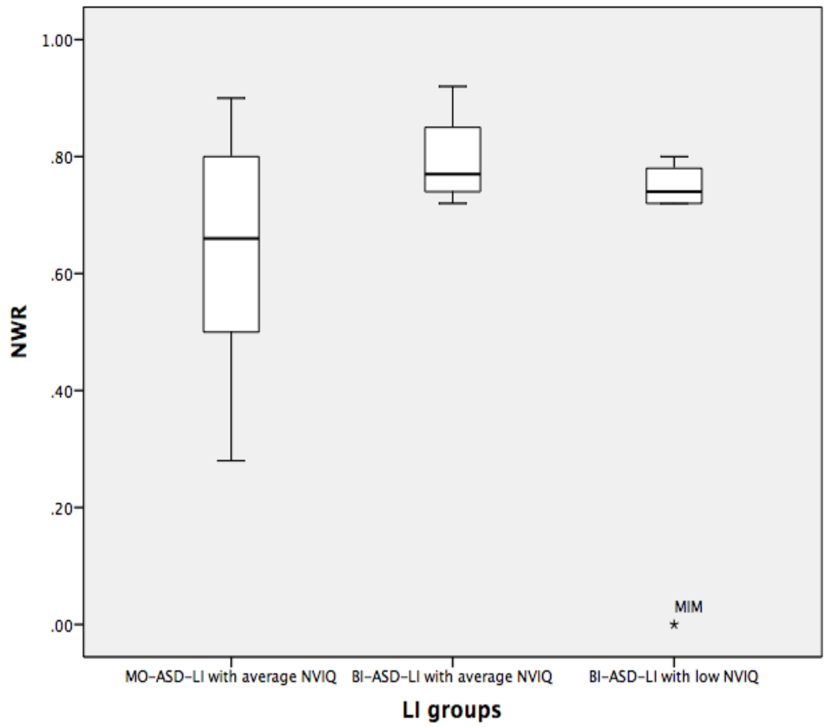
More than half (16/30) of ROB's production was characterised by inappropriate answers: sometimes he did not repeat the sentences (*n* = 8), for others he selectively repeated some lexical items of the sentence, e.g. stimulus: 'the parents put away the toys', ROB's production: 'the toys' (*n* = 4) or he commented on the sentence without repeating it, e.g. stimulus: 'the rabbit wants to eat the salad now', ROB's production: 'Yes, he is hungry' (*n* = 2) and finally instead of repeating the object wh-questions given as stimuli he answered the questions, e.g. stimulus 'Which boy is the man drawing?', ROB's production: 'He is drawing a robot' (n = 2). Similarly AFP behaved unusually on 1/3 of the task (11/30 sentences): in eight occurrences he repeated only the second half of the sentences (independently of the computational complexity of the sentence), while in three cases he changed the lexical items of the sentences with others not included in the task (e.g. stimulus: 'the dad can drive the car very well', AFP's production 'the dad he can drive the vehicle').

Finally MIM was the only case (in all our entire ASD population sample) for which 80% of the answers were unintellegible (24/30), leading to an impossibility of transcribing and coding his sentences.

The kinds of behaviours showed by the ASD-LI with low NVIQ group were not found in any other profile of bilingual children with ASD nor in any children of the monolingual group with ASD, except for the two monolingual children YAT and FIZ who had a LI profile with low NVIQ. As explained in section 6.3.3, YAT performed very badly in the second half of the test, while FIZ showed a high rates of no responses (20% of the task), essentially due to inattention during the testing phase.

Concerning phonological abilities, seven children showed impaired performance on the NWR task: two of them CAT, RIV) were included in the ASD-LI with average NVIQ profile, while five (ROB, AFP, MVI, MIM, YVA) were included in the ASD-LI with low NVIQ profile. Figure 58 shows general group performance for the global score of correct repetition on the NWR task.

FIGURE 58. MEAN RATE OF CORRECT REPETITION FOR NWR IN LI GROUPS



No significant difference was found between the MO-ASD-LI with average NVIQ group and the BI-ASD-LI with average NVIQ group ($U(17) = 12$, $p = .099$), between the BI-ASD-LI with low NVIQ group and the MO-ASD-LI group ($U(18) = 19$, $p = .610$), and between the two BI-ASD-LI groups ($U(8) = 7$, $p = .459$). Similarly, no differences were found between the three groups on syllable length and on number of consonant clusters, as shown in Table 47. All three groups showed a significant drop in performance on three-syllable items and on nonwords containing two consonant clusters, indicating that bilingual children with language impairment were equally sensitive to phonological complexity as their monolingual peers. The only *outlier* was MIM whose speech, as in the SR task, was so severely disordered that he was unintelligible, yielding to the impossibility of a transcription and coding of his answers.

TABLE 47. MANN-WHITNEY INTERGROUP COMPARISONS ON SYLLABLE LENGTH AND CONSONANT CLUSTERS

		1-syll	2-syll	3-syll	0-clus	1-clus	2-clus
MO-ASD-LI/BI-ASD-LI with average NVIQ	*U*	6	6	11	4.5	3	9.5
	p	.271	.270	.855	.170	.098	.603
MO-ASD-LI/BI-ASD-LI with low NVIQ	*U*	9	16	14.5	12.5	29	24
	p	.056	.138	.101	.065	.916	.472
BI-ASD-LI with average NVIQ/BI-ASD-LI with low NVIQ	*U*	5	5	4.5	4.5	3	.3
	p	1.00	1.00	.839	.839	.421	.388

To sum up:

Two profiles of language impairment were detected in the BI-ASD group, a BI-ASD-LI with average NVIQ profile and a BI-ASD-LI with low NVIQ profile. We compared the performance of both these groups with the ASD-LI profile previously detected in our monolingual group of children with autism, in order to see whether bilingualism would affect the performance of children with ASD on tasks of SR and NWR. Results showed that the children with BI-ASD-LI with average NVIQ profile resembled their MO-ASD-LI peers both on morphosyntax and phonology. The children in the BI-ASD-LI with low NVIQ group, instead, showed a profile similar to the other two LI groups on phonological abilities only. On the SR task, this group showed a very unique behaviour that was characterised by uneven performance (discrepancies of performance between the first half and the second half of the task, presumably due to selective drop of attention) or unusual responses (e.g.

answers to the questions, comments on the sentences, etc). Moreover, for one child in the ASD-LI with low NVIQ profile, MIM, it was impossible to transcribe and code his productions because his speech was so severely disordered that he was unintelligible.

Before moving to the final conclusions, we checked for possible relation between bilingualism and structural language / NV abilities in order to verify that knowing a second language did not constitute a source of either positive or negative biases.

9.4. Do bilingual factors predict children's performance on LITMUS tasks and NV cognitive tasks?

We looked at the relation between bilingual factors previously selected in section 8.5.3.1 (AoO, LoE, early language exposure, current language use at home, current language richness, French language dominance, and PLD) and performance on SR (% identical repetition) and NWR (% correct repetition) for language abilities, and performance on RPM, Block Design and Matrix Reasoning for NV cognitive level. MIM was excluded from the correlation on language abilities because both his performance on SR and NWR were phonologically unintelligible, so we could not properly estimate his linguistic skills. Table 48 shows that no significant correlations were observed between bilingual factors and scores on either SR and NWR or on NV tasks, indicating that bilingualism did not play a significant role in the performance of our group of children with BI-ASD.

TABLE 48. BIVARIATE PEARSON CORRELATIONS BETWEEN BILINGUAL FACTORS AND SCORES ON SR AND NWR TASKS

Bilingual factors (PABIQ)	% SR identical repetition	% NWR correct repetition	RPM (standard scores)	Block Design (standard scores)	Matrix Reasoning (standard scores)
AoO French	.077	-.077	.355	.291	.414
LoE French	.285	.200	.087	.291	.414
Early L1 exposure	.045	.499	.142	.190	.203
Early French exposure	-.096	-.065	.149	.193	.202
Current L1 use at home	-.059	.218	.052	-.420	-.208

Current French use at home	.077	.123	-.070	.462	.265
Current L1 richness	.008	-.081	.104	-.232	.038
Current French richness	-.067	-.326	.102	.357	.003
French Language Dominance	-.006	-.232	-.066	.321	.155
PLD	.038	.030	.167	.159	.165

9.5 Conclusions and discussion

The main objective of this chapter was to investigate whether bilingual children would display different structural language/NV ability profiles from monolingual children with ASD. The results of our pilot study, conducted on a small group of 14 BI-ASD children aged 6- to 12- years taken from the whole spectrum, suggested that bilingual children performed in line with their monolingual peers showing no significant differences on tasks evaluating structural language (in line with Baldimtsi et al., 2016) and NV cognitive abilities (in line with Baldimtsi et al., 2016; Gonzales- Barreiro & Nadig, 2017, and more generally with studies on TD bilingual children: Bialystok & Martin, 2004; Hilchey & Klein 2011; a.o.). The structural language/NV ability profiles found in our bilingual children with autism perfectly matched the (four) profiles found in our MO-ASD group (Chapter 6), leading to the conclusion that bilingualism was not a detrimental factor to children with autism (contra Baron-Cohen & Staunton, 1994), as it has already been proved for children with SLI and children with Down's syndrome (see section 8.2). Crucially, the existence of a double dissociation of language and NV abilities like the one found in the monolingual group was also found in the bilingual children (in line with Smith & Tsimpli, 1995). Furthermore, the existence of a fifth profile, the ASD-LI with low NVIQ profile, which we hypothesised also for monolingual children although it did not clearly emerged from our cluster analyses, was found in the BI-ASD group. The fact that a higher rate of intellectually impaired children was included (by chance) in the bilingual group (16 % in the MO-ASD group vs. 42% in the BI-ASD group) enabled the ASD-LI with low NVIQ profile to clearly emerge as a profile *per se*. This also allowed the two monolingual children with LI and low NVIQ to join this profile.

Similarly to what has been found in studies on bilingual children with SLI (Armon-Lotem & Meir, 2016; de Almeida et al., 2017; Marinis et al., 2017; Tuller et al., 2018; a.o.), both LITMUS tasks proved to be reliable tools for both disentangling the effects of bilingualism from those of language impairment in autism, yielding to a clear identification of LI profile among bilingual children, and pinpointing the sources of impairment in phonology and morphosyntax (difficulties with computational complexity).

The fact that the bilingual children distributed into the existing profiles of monolingual children with ASD raised the question as to what extent both profiles of children with BI-ASD-LI were similar to the profile of language impairment found in the MO-ASD group and the three LN profiles in the BI-ASD group were reminiscent of the three LN profiles in the monolingual children with autism in quantitative and qualitative perspectives. The analyses on SR and NWR showed that the language profiles of the bilingual children with ASD-LI and average NVIQ resembled the LI profile in the MO-ASD group both for morphosyntactic and phonological abilities, in terms of computational complexity effects and error typology. The same was true of the three LN profiles in the BI-ASD and MO-ASD groups. In contrast, the bilingual children with ASD-LI and low NVIQ displayed some unusual behaviours on the task of SR (performance on the NWR was similar to the MO-ASD-LI group). These unusual behaviours included not only selective drops of performance due to lack of attention during task administration but also production of inappropriate answers (responses to questions, comments on the sentences, etc.), which were reminiscent of findings by Prévost et al. (2017) and Zebib et al. (2013) in their studies on production of wh-question in French-speaking children with ASD (section 1.2.6.2.2). It is very important to notice that two monolingual children with ASD-LI and low NVIQ behaved the exact same way as their bilingual peers. This fact further proved that bilingualism did not affect language and cognitive abilities of our group of children. Moreover, among the BI-ASD-LI with low NVIQ group, one child was excluded from analyses because of his unintelligible speech, which made transcription of his production impossible. This case was reminiscent of the one reported by Wolk & Edwards, (1993), who concluded for the impossibility of evaluating language abilities in an 8-year-old boy with autism with severe articulation disorders.

The unusual behaviours limited to the SR task and to the specific ASD-LI with low NVIQ profile and the presence of cases like the child with severe articulation disorders raise the following questions: (1) to what extent can morphosyntactic abilities be evaluated via SR in children with LI and low NV abilities? (2) to what extent may the behaviour of these children be due to the coexistence of both LI and low NV abilities? and (3) to what extent is it possible to assess structural language abilities of children that cannot produce intelligible speech and more generally of children that do not talk at all (recall that nonverbal children constitute the 30% of the children on the spectrum)? We will expand and speculate on these topics in the discussions of the present work (Part IV).

In conclusion, this was the first study that looked at structural language/NV ability profiles in bilingual children with ASD, comparing their performance with monolingual children with ASD. Bilingualism did not affect performance of children with ASD. Further evidence for this was the fact that bilingual factors, such as LoE, AoO, language richness and PLD, which have often been pinpointed as good indicators of language outcomes in bilingual children with and without SLI (see Tuller et al., 2018), were not related to either language abilities (as measured by SR and NWR) or NV scores on fluid reasoning and visuospatial tasks. Nonetheless, these results on bilingual children with ASD need to be replicated in larger samples. Moreover, it may be the case that the fact our sample included only simultaneous bilinguals and that most children were French-dominant led to little heterogeneity in ability profiles. It would be interesting to replicate the same experiment on a sample including L1 dominant bilingual children with ASD.

Part IV

General discussion

Autism Spectrum Disorder is characterised by notable heterogeneity of phenotypic profiles of both language and cognitive abilities. The main aim of this study was to find a methodological strategy that allows for the integration of both of these features (evaluated via the most reliable measures possible) in order to **identify** (prove their existence) and **describe** (define their phenotypical realization) profiles of structural language and NV abilities in monolingual and bilingual children with ASD.

In this section we discuss the main results of our work, integrating results reported in previous studies and focusing on possible theoretical implications and directions for future research. For a detailed analysis of all the results we presented in the present work we refer the reader to the discussion and the conclusion of each one of the preceding chapters.

The methodological procedure behind the **identification of language/cognitive profiles** was inspired by the findings reported in Joseph et al. (2002), Kjelgaard & Tager-Flusberg (2001) and Tuller et al. (2017), which, via inclusion of children with autism along the whole spectrum, demonstrated that it is possible to document the existence of (at least) four different phenotypical profiles derived from the combination of language and cognitive abilities in verbal children with autism: ASD-LN with normal IQ, ASD-LI with normal IQ (often compared to SLI), ASD-LI with low IQ, and ASD-LN with low IQ (the least attested profile). The fact that only three (out of 16) studies that investigated language abilities in monolingual children with ASD by including both HF and LF children (see Chapter 1) detected the existence of all four profiles raised the question as why their existence has not been attested more often in the literature. Assuming that these four profiles do exist, we hypothesized that the answer to the question was twofold. (1) While the linguistic abilities of children with ASD and normal IQ have often been investigated, the relative infrequency of studies that have included children with low cognitive profiles in their population samples and the low number of children with low cognitive abilities included in these same studies, have led to a lack of knowledge about the capabilities of these children, especially of children with a ASD-LN with low IQ profile (8% of participants in Joseph at al., 2002; 10% in Kjelgaard and Tager-Flusberg, 2001, and 5% in Tuller et al., 2017). (2) The literature on both linguistic and cognitive abilities in children with ASD lacked in consistency and appropriateness regarding how both linguistic and intellectual capacities have been tested. Language abilities have typically been assessed through vocabulary testing or developmental scales, an approach that have circumscribed the investigation of language in autism to vocabulary, neglecting underlying structural language abilities (morphosyntax and phonology). As to cognitive abilities, they have often been assessed through FSIQ scores, which have been reported to be hard to interpret because of the great heterogeneity of autistic abilities across different intellectual domains (nonverbal, verbal, working memory, speed processing). The appropriateness of these scores has also been questioned for the investigation of the relationship between intellectual and linguistic profiles, in contrast to nonverbal measures.

We reasoned that addressing methodological shortcomings (1) and (2) would make it possible to distribute verbal children with ASD into homogeneous subgroups of abilities, and it would lead to replication (and maybe expansion) of findings reported by Joseph et al. (2002), Kjelgaard & Tager-Flusberg (2001), and Tuller et al. (2017).

In our study we addressed the first point (1) by including children from the whole spectrum, specifically aiming for the inclusion of both children with normal and low cognitive abilities. For the second point (2), we identified and justified use of measures most likely to appropriately assess structural language and nonverbal cognitive abilities in ASD.

Moreover, we hypothesised that the same four profiles, whose existence was evoked for monolingual children with ASD, could potentially be found in bilingual children with ASD as well. This hypothesis was mainly based on the very few studies that investigated structural language and NV abilities in bilingual children with ASD, notably Baldimtsi et al. (2016) and Smith and Tsimpli (1995), and on previous findings on other bilingual clinical populations. Notably, we built up our prediction on studies that demonstrated that bilingual exposure did not cause additional language delays for children with SLI (Armon-Lotem & Meir, 2016; de Almeida et al., 2017; Marinis et al., 2017; Tuller et al., 2018 a.o.) and children with Down Syndrome (Feltmate & Rehab, 2008), and on studies that showed that bilingual children did not underperform monolingual children on NV cognitive tasks (Bialystok & Martin, 2004; Hilchey & Klein, 2011; Jarvis et al., 1995). We thus assumed that bilingualism would have neither detrimental nor enhancing effects on children with ASD, and that bilingual children with ASD would behave like monolingual children with ASD, showing the same profiles of structural language / NV abilities. For these reasons we chose measures of structural language abilities and nonverbal cognitive abilities that could be used not only for evaluating monolingual children but also bilingual populations.

With all these assumptions in mind, the first step of our work sought to identify the best tools for evaluating structural language/NV abilities in children with ASD. We first ran our analysis on monolingual children, including 37 children with ASD aged 6- to 12- years from the whole spectrum. We then integrated into this analysis a small group of 14 bilingual children with ASD aged 6 to 12 years, which we compared with the monolingual children with ASD.

Concerning the choice of language measures, we followed the hypotheses of Wittke et al. (2017) that standardized tasks may be inadapted for evaluating detailed morphosyntactic and phonological abilities of children with ASD and of Prévost et al. (2017) which suggested that experimental tasks constructed for investigation of specific aspects of morphosyntax and phonology may obscure the real underlying capacities of children with ASD, requiring functional communication skills and competence in social interaction (e.g. act-out, sentence-picture matching, elicited tasks). We argued for the use of two specific repetition tasks, LITMUS-SR and LITMUS-NWR. We hypothesised, and showed, by comparing performance on standardized tests and the LITMUS tasks, that SR and NWR were the best suited measures for evaluating morphosyntactic and phonological abilities respectively, in both monolingual and bilingual children with ASD (Chapter 5 and Chapter 8). These tasks had already been proved to be sensitive for the detection of SLI in monolingual and bilingual children. They were specifically created to target particular aspects of computational complexity, minimizing the intervention of other aspects (pragmatic difficulties, working memory effects, etc.) that can mask the "real" capacities of children with ASD.

Concerning cognitive abilities we followed previous studies by Barbeau et al. (2013), Housekeeper (2011), Nader et al. (2016), Stevenson & Gernsbacher (2013), a.o., which argued for the use of nonverbal tasks, such as Raven's Progressive Matrices (RPM), and two subtests of the Perceptual Reasoning Index of the 4th edition of the Wechsler Intelligence Scale for Children (WISC-IV), Block Design and Matrix Reasoning, as the best-suited measures for evaluating cognitive abilities in children with ASD. We justified our hypothesis by comparing performance of both monolingual and bilingual children with ASD on these three NV tests with other measures of cognitive abilities, notably FSIQ scores, Verbal Index scores, etc. The NV tests proved to be the most adapted because they could be easily used on children from the whole spectrum and they were minimally linked to the verbal abilities of our participants, providing the foundations for meaningful exploration of the relation between linguistic abilities and cognitive abilities in ASD (Chapter 5). Moreover, they were proved to be unrelated to the nature of the children's language exposure, since the monolingual and bilingual children with ASD displayed the same performance on all tasks (Chapter 8).

After having explicitly argued for the use of specific measures of formal language abilities (SR and NWR) and of NV cognitive abilities (RPM, Block Design and Matrix Reasoning), we sought to explore the existence and the main characteristics of profiles of structural language and nonverbal abilities obtained crossing both of these abilities. We argued for the use of an unsupervised machine learning approach, notably the cluster analysis, in order to limit as much as possible the number of *a priori* assumptions about our profiles' subtyping. Indeed, we wanted the main characteristics used to group the children into profiles of abilities to be as little controlled as possible. Cluster analyses were first run on the group of monolingual children with ASD (n = 37), resulting in four clusters for both morphosyntactic / NV abilities and phonological / NV abilities. These analyses confirmed the existence of three profiles previously detected in the literature, ASD-LN with high NVIQ, ASD-LI with average NVIQ, and crucially, ASD-LN with low NVIQ. A fourth profile, ASD-LN with average NVIQ, was detected as well, a result not previously alluded to in the literature. From the analysis of the monolingual children, no ASD-LI with low NVIQ profile clearly emerged, although two children in the MO-ASD group selectively displayed both impaired structural language and cognitive abilities. They were too few, however, to constitute a separate profile. The existence of this fifth, and expected, profile emerged only when the bilingual children with ASD were integrated into the analysis. This led to the inclusion of a somewhat large number of intellectually impaired children in the analysis (12/51 children with low NVIQ), which in turn enabled the ASD-LI with low NVIQ profile to clearly emerge separately (including the two monolingual children with LI and low NVIQ).

In conclusion, our analysis demonstrated that with the use of proper tools and the inclusion of a minimum number of children from different parts of the whole spectrum, all four profiles that had been previously evoked in the literature could be identified in both monolingual and bilingual children with ASD. In addition, our analysis detected the existence of another profile, ASD-LN with average NVIQ. The existence of three distinct profiles among LN children and two distinct profiles among LI children suggest that it is inadequate to divide children simply on the basis of their language abilities (ASD-LI and ASD-LN) without considering the possible relations with cognitive capacities. Proceeding in such a way results in having groups that might obscure multiple profiles of abilities and consequently jeopardise our understanding of the real capacities of children with ASD.

However, the identification of the profiles was not sufficient to understand their true nature. After the cluster analysis, we still did not know how grammatical abilities may differ across these subgroups and whether cognitive abilities may have played a role in differentiating children with LI and average or low NVIQ from children with LN and high, average or low NVIQ.

We aimed to answer these issues by investigating the nature of the five structural language/NV ability profiles we detected in our analyses. Concerning **the description of** the five **profiles**, we mainly focused on the phenotypical realization of the children's structural language abilities, comparing their performance on SR and NWR with control populations. There have been several controversies, especially in morphosyntax, concerning the authenticity of parallels drawn between ASD-LI and SLI, and between ASD-LN and TD children (Chapter 1), with some studies indicating that from the perspective of a qualitative error analysis children with SLI are more impaired than children with ASD-LI ((Riches et al., 2010; Roberts et al., 2004; Prévost et al., 2017) and that children with ASD-LN show language abilities which are not spared compared with their TD age peers (Durrleman et al., 2017; Modyanova et al., 2017; Terzi et al., 2017; Tuller et al., 2017). We sought to verify whether the division of children into profiles of language/NV abilities, like the ones obtained in our first analysis, would help clarifying the issue.

We compared the performance of children from the three ASD-LN profiles with two groups of TD children, one group of younger TD children (*n* = 42) aged 4 to 5 years and a group of age-matched TD children (*n* = 42). The children with ASD-LI were compared with 26 age-matched children with SLI. Comparative analyses were run on group results and individual results from quantitative, qualitative and developmental perspectives. Results of our analysis showed that while the phenotypical realisation of phonological abilities in the three ASD-LN profiles did not differ from that of younger TD children, for morphosyntax the picture was more complicated. Our results seem to suggest that there is a subgroup of children with an ASD-LN profile (ASD-LN with high NVIQ) that displays "normal" structural language abilities, and two subgroups of children with ASD-LN (ASD-LN with average NVIQ and the ASD-LN with low NVIQ) that show some selective drop in performance on structures involving highly demanding derivations (computationally speaking), as in the case of relative clauses. Crucially, our results seem to clarify previous findings in the literature suggesting a possible explanation for the fact that some studies found age-matched performance in the ASD-LN profile while others did not, in line with Baldimtsi et al. (2016), Durrleman et al. (2017), Modyanova et al. (2017), Terzi et al. (2017), and Tuller et al. (2017). On the basis of our results, there seem to be two phenotypical realisations of language abilities in children with ASD-LN: one that displays structural language abilities similar to TD age-matched children (displayed by children with high NVIQ) and another whose syntactic abilities are not completely spared.

Concerning the supposed similarities between the ASD-LI profile(s) and SLI, mixed results were obtained. While the phenotypical realisation of phonological abilities in both the ASD-LI with average NVIQ profile and the ASD-LI with low NVIQ profile resembled what is found in SLI, the three groups differed on morphosyntactic abilities. Performance on SR suggested that both groups of children with ASD-LI were more severely impaired than the children with SLI (in line with Loucas et al., 2008; Modyanova et al., 2017; Sukenik & Friedmann, 2018; Taylor et al., 2014) on all measures, which included group results and individual results on Identical repetition and Target structure score, on the whole task as well as on each substructure (less complex and more complex conditions). In addition, qualitative error analysis confirmed not only the phenotypical difference between the children with ASD-LI and the children with SLI, but also some dissimilarities between the children with ASD-LI and average NVIQ and the children with ASD-LI and low NVIQ. While the children with SLI tended to produce wholesale changes to the structures of the sentences, diminishing computational complexity by simplifying more complex structures (e.g. transforming object relatives into subject relatives, finite argument into declaratives, etc.) (in line with Riches et al., 2010), the children with ASD-LI with average NVIQ were more prone to produce omission errors, which could range from simple determiner omission to omission of one or several fundamental elements of the sentences. Crucially, while the behaviour of the children with SLI was limited to more complex conditions (wh-object questions, argument clauses and relative clauses), the children with ASD-LI produced omission on all structures and conditions, suggesting that their impairment was much more severe and pervasive than the one found in SLI. In addition, the children with ASD-LI and low NVIQ differed from both the children with ASD-LI and average NVIQ and the children with SLI. These children displayed unusual behaviour on the tasks, including selective drop of performance on half of the stimuli and production of inappropriate answers. These results were reminiscent of the findings in Prévost et al. (2017) and Zebib et al. (2013) which highlight the possible production of unexpected answers by children with ASD. These results seem to indicate that different mechanisms might be in place in the three groups and that that at least the LITMUS-SR task is not completely immune to the impact of certain factors, e.g. pragmatic impairment.

The analysis of the phenotypical realisation of language abilities in the three LI profiles showed that both groups of children with ASD-LI displayed much more severe impairment in morphosyntax than the children with SLI, while their impairment did not differ on phonological abilities. These results question the legitimacy of the hypothesis that there is an SLI profile in ASD (in line with Bishop 2003, 2006, 2010; Conti-Ramsden et al., 2006; Tomblin, 2011; Whitehouse et al., 2007; Williams et al., 2008; a.o.). This conclusion strengthens our decision to adopt the label of ASD-LI, intended as a neutral label which does not carry any reference to an SLI profile for children with ASD within the interpretation of the two conditions as part of a continuum.

Regarding the possible influence of bilingualism on the performance of children with ASD, our results demonstrated that the tools used for evaluating structural language abilities and nonverbal cognitive abilities in monolingual children with ASD were well-suited for evaluating bilinguals as well, in line with previous research on other clinical populations, such as SLI (Armon-Lotem & Meir, 2016; de Almeida et al., 2017; Marinis et al., 2017; Tuller et al., 2018 a.o). Moreover, the same profiles of abilities as the ones found in the monolingual children with ASD were found in the children with exposure to a second language, indicating that bilingualism is not a detrimental factor for children with ASD, similarly to what have been demonstrated in other studies on bilingual children with ASD (Baldimtsi et al., 2016 and Smith & Tsimpli, 1995) and for other clinical populations, notably SLI (Altman et al., 2016; Armon-Lotem & Meir, 2016; Grimm & Schulz, 2014; Marinis et al., 2017; Paradis, 2007; Tsimpli et al., 2016; Tuller et al., 2015) and Down Syndrome (Feltmate & Rehab, 2008).

We have thus far summarized the main results of our work, integrating previous findings from the literature. We now discuss how our results speak to theoretical debates regarding three principle topics: the existence of an SLI profile in children with ASD, the relation between cognition and language and, finally, bilingualism in ASD.

Is there a phenotypical SLI profile in children with ASD?

The fact that both groups of children with ASD-LI and children with SLI showed similar performance on NWR, while they behaved differently on the SR task inevitably raises the following question: to what extent can we talk about different profiles of abilities between ASD-LI and SLI? Our results seem to suggest that children with ASD-LI show the same phenotypical profile as children with SLI for phonological abilities, while they display a different profile for morphosyntactic skills. Before jumping to quick conclusions, we think that some considerations should be made.

One possible reason for the different results on the two tasks could be methodological. It could be the case that the analysis we ran on NWR was not sufficiently detailed to detect possible underlying differences between the ASD-LI and the SLI groups. Running a much more extensive qualitative error analysis on the NWR task, similar to the one we ran for SR, could result in children with ASD-LI displaying different errors from the ones made by children with SLI. This would lead to the conclusion that the underlying characteristics of language impairment in children with ASD are different from what is found in SLI (in line with previous works by Bishop et al., 2010 and Tomblin et al., 2011).

A second possible reason for the different results between NWR and SR could be the type of impairment displayed by the children with SLI. The results from the older children in our SLI group on SR were unexpected. Often children with SLI do not show significant improvement in their syntactic development after childhood (see Riches et al., 2010). It could be the case that these older children were less impaired than the older children with ASD-LI on morphosyntax, which could have led to some significant differences in the results from SR. However, this hypotheis does not explain why the two populations make different types of errors. In a similar vein, a difference in length and type of speech language therapy, more targeted on morphosyntactic impairment for children with SLI, could have resulted in older children with SLI performing better (see section 7.3.4). For children with ASD, language impairment is one among a kaleidoscopic inventory of deficits, so we can suppose that the time spent focusing on morphosyntax is less than for children with SLI.

Third, it could be that processing and repeating sentences is more complex than processing and repeating nonwords. Repeating a sentence involves not only processing all of its syntactic components, it also involves processing its articulation. Moreover, producing a sentence needs planification of syntactic and semantics elements, which constitutes an additional step with respect to comprehension. The children with ASD-LI may have had difficulty integrating the syntactic and semantic properties of the sentences in the SR task, bringing their performance down, in contrast to the children with SLI. The language impairment phenotype then would be limited to deficits in syntactic complexity for children with SLI (failure to process complex structures and production of simpler sentences instead), while in children with ASD-LI the high rate of omission of parts of the sentence could be the phenotypical realisation of an impossibility to process several types of information at the same time. This hypothesis may also explain why the children with ASD-LI and the children with SLI did not differ on phonological abilities. In NWR the process of integrating multiple types of information is reduced, since nonwords do not carry any semantic value. If this hypothesis is correct, NWR could effectively be considered as a pure task of phonological abilities. For evaluating morphosyntax, instead, along with the SR task which has been proved to be a good detector of expressive morphosyntactic abilities in ASD population, we might need to evaluate children with ASD also via carefully constructed comprehension tasks.

Although we have tried to “purify” as far as we could the measures of syntactic abilities in children with ASD, using an SR task specifically created to concentrate on computational complexity, the difference in performance between SR and NWR raises the question as to what extent the linguistic competence (in a Chomskyan sense) of children with

ASD-LI was tapped exclusively by the use of SR. We might suppose that children with ASD-LI do have the same profile of language abilities as children with SLI, including morphosyntax, but that SR is still not enough by itself to measure their real underlying morphosyntactic abilities. However, while it is interesting to speculate on the origins of difficulties in sentence repetition we still know little about the cognitive mechanisms underpinning this task. For example, the extent to which repetition depends on successful comprehension is not known, and therefore we cannot necessarily assume that the phenotypical realisation of error types (mainly omission) found in the ASD-LI with average NVIQ group and the unusual behaviours found in the ASD-LI with low NVIQ group were necessarily the mirror of underlying language impairment in ASD. It could be the case that children with ASD-LI comprehend sentences at the same level as their SLI peers, but that different mechanisms may be at play when processing for production occurs. If it is the case, what other ways of evaluating morphosyntactic abilities of children with ASD could there be? On-line measures of comprehension, e.g. looking-while-listening tasks, could be one possible indicator of further information on morphosyntactic abilties. We think, in fact, that both production and comprehension should be evaluated together in order to have a clearer picture of the real capacities of children with autism. While for production we may have found the most adapted tools, for comprehension some work needs to be done. In this vein, we think that the choice of comprehension tasks that have normally been used with the ASD population, would cause the same problems as the available standardized production tasks (as seen in Chapter 5). The idea would be to build a comprehension task, designed to be a "minimalist" assessment of morphosyntax (similarly to what we did for production), which would include a variety of structures of different complexity and that would answer all of carefully argued critieria. A few studies have already shown that use of eye-tracking techniques may lead to fine-grade measures of language abilities in children with ASD (see Chita-Tegmark et al., 2015; Léger, 2017; Naigles & Fein, 2017). The use of eye-gaze measures could offer some important insight about the underlying causes of the possible deficits in syntactic abilities that production tasks (including SR) may flag. First, by lowering the response demand and minimizing the communication demand, eye-gaze measures can rule out the possibility that lower scores on syntactic abilities are simply by-products of production demands. Second, by measuring language processing in real time this type of task can determine whether language deficits are caused by poor processing efficiency itself (a deficit in processing during comprehension) or are instead related to other factors such as social and communication problems (a deficit in processing during production). Moreover, passive evaluation of linguistic competence could provide further information on morphosyntactic abilities of children with ASD-LI and low NVIQ, circumscribing the possible effects of unusual behaviours that could be rather observed andor specifically related to production. Finally, the use of eye-tracking techniques, being completely passive, could be extended to children with severe articulation problems (such as the child that was excluded from our results in the bilingual study) and even to minimally-verbal/nonverbal children with autism (see Cantiani et al., 2016).

Cognition and language: which is the relation in ASD?

The other crucial outstanding issue of this work was the relation between cognitive abilities and structural language capacities in children with ASD. From the vantage point of the existence of a language module in the human mind/brain, which thus can be selectively spared (see Smith & Tsimpli, 1995), our results receive a natural interpretation. The existence of a double dissociation like the one found in the ASD-LI with average NVIQ profile and the ASD-LN with low NVIQ profile in both monolingual and bilingual children with ASD indicates that children with autism can indeed display impaired language abilities in presence of spared nonverbal intelligence or spared language abilities in the presence of impaired nonverbal intelligence, a profile reminiscent of that found in Williams Syndrome (Mervis & Velleman 2011) and also in the language Savant *Christopher* (Smith & Tsimpli 1995).

Specifically, concerning the existence of the ASD-LN with low NVIQ profile in both monolingual and bilingual children, our results support the few studies that have reported its existence in the literature. Among all of the children displaying low NVIQ (12/51 children, roughly 23% of our total population, pulling together monolingual and bilingual participants), 7 children displayed a LN profile for morphosyntax (n = 1), phonology (n = 2) or both (n = 4). We note that the frequency of this profile, whose existence was only found in very few studies, constituted 13% of the total sample, which is slightly higher than what Joseph et al. (2002), Kjelgaard & Tager-Flusberg (2001) and Tuller et al. (2017) found (respectively 8%, 10% and 5%). Crucially, children with an ASD-LN with low NVIQ profile constituted more than half of the children with low NVIQ present in our sample (7/15), while in other studies their number was generally lower (Joseph et al.: 6/32; Kjelgaard & Tager-Flusberg: 11/31; Tuller et al.: 1/6). Including children with intellectual deficiency in the evaluation of language abilities is fundamental for a better understanding of the “real” capacities of children on the autism spectrum. For many years, the assumption that low cognitive abilities must entail low language abilities has been taken for granted. However, these results, along with the existence of profiles like the ones found in SLI and Williams Syndrome, reinforce the idea that language may constitute an independent module in the brain (see Chomsky, 1980; Fodor, 1985; Smith & Tsimpli, 1995; a.o.), even though this module interfaces with other modules and central systems. We can further speculate on this interaction for partially explaining the unusual behaviours displayed by the children with ASD-LI and low NVIQ. We can hypothesise that the co-occurrence of both language impairment and low nonverbal cognitive abilities may further pull down performance of children with ASD (phenotypically realised by unusual behaviours on the SR task). Similarly, we can suppose that the co-occurrence of normal language abilities and high NVIQ may pull up the performance of children with ASD, resulting in a profile that displays the same structural language abilities as TD children. However, although these two particular profiles seem to indicate that some sort of interaction takes place between language and cognition, we cannot conclude which domain influences which or whether there is mutual influence.

Bilingualism and ASD

The results of our pilot study on 14 bilingual children with ASD suggest that children with ASD do not display any detrimental factors from being bilingual (contra Baron-Cohen & Stauton, 1994). The presence of children in all five profiles of structural language/NV abilities indicated that bilingual children with ASD show the same capacities as their monolingual peers both for structural language and NV abilities (in line with Hambly & Fombonne, 2012; Kay-Raining Bird et al., 2012; Ohashi et al., 2012; Petersen et al., 2012, Valicenti-McDermott et al., 2013). However, our results did not find any beneficial effect in bilingualism either (contra Gonzales-Barrero & Nadig, 2017). These results partially confirm findings by Baldimtsi et al. (2016) which revealed no differences between monolingual and bilingual populations on syntactic complexity at a microstructural level, but which reported a beneficial effect of bilingualism for story structure complexity and avoidance of the use of referentially ambiguous forms. These findings are crucial in making cliniclians' recommendations to parents of bilingual children with ASD positive to continue speaking their home language to their children, independently of their cognitive level (contra Baron-Cohen and Staunton, 1994). The existence of children presenting the ASD-LN with low NVIQ profile is an indicator that language abilities can be spared in presence of low cognitive abilities in the bilingual population with ASD as well (in line with Smith & Tsimpli, 1995).

Limitations, future directions of research and clinical implications

Additional research is needed to test the replicability of the current findings and to address the following limitations. The number of participants in the current study was limited; a larger population sample should make it possible to address the question of the relative prevalence of the profiles, notably the discrepant ASD-LN with low NVIQ profile, especially if we hope for the general use of unsupervised classifications, such as cluster analysis. Running the same analysis on a much bigger cohort of children with ASD may further detail the profiles of structural language/NV abilities and even increase the number of profiles detected in the spectrum.

A more targeted analysis of NWR via in-depth error analysis could help clarify the existence of an SLI profile in ASD. Moreover, the use of passive tasks, such as eye-tracking tasks, might both provide new means for tapping the underlying capabilities of children with ASD and at the same time give researchers the opportunity to include minimally-verbal and nonverbal children with ASD in studies focusing on language abilities.

Moreover, the pilot study on bilingual children with ASD should be replicated with a higher number of participants, crucially including L1 dominant bilinguals, in order to verify whether the profiles would be affected by a more heterogeneous population. Furthermore, our study on bilingual children lacked a comparison with bilingual control populations. Further research should compare performance of BI-ASD children with BI-SLI and BI-TD groups.

Concerning the ASD-LN with low NVIQ profile, future research should include a direct comparison with children with Williams Syndrome to see whether the phenotypical realisations of both these conditions are different or not, similarly to what has been done for the ASD-LI and SLI profiles.

Regarding the identification of profiles of abilities in ASD, we think that several pieces are still missing from the puzzle. Our work has proposed a step forward into identifying structural language / NV ability profiles in children with autism, but other factors, that we did not have the possibility of investigating in the present work, should be taken into consideration in future analyses. The integration of other factors (evaluated via the most reliable measures possible) should further specify the types of profiles present in autism and help develop treatments tailored to the specific needs of individuals, based on their particular patterns of strengths and impairment. We argued, for example, that severity of autism symptoms was not a detrimental factor, either for structural language or for NV cognitive abilities. However, it could be the case that the scales we used as measures of autism severity were not pure enough to be put in relation with language and NV abilities. The CARS, the ECA-R and the ADOS contain several items evaluating linguistic and cognitive skills. We think that future research should extend the methodology used in the present study for language and cognition to autism severity. A first step in this direction has been made by Georgiades et al., (2013) and Wiggins et al., (2012), which via unsupervised machine learning analyses have already identified clusters of autism severity in ASD. Similarly, future research should also refine our analysis by including measures of Executive Functions, such as tasks evaluating attention and inhibition (possibly as nonverbal as possible). A careful choice of these measures could particularly help better define the impairment displayed by children with ASD-LI and low NVIQ, who showed a selective drop in performance on SR, maybe because of attention deficits.

Finally, we should not forget that what we pictured here was the profile of children with ASD at a given time of their developmental trajectories. We suggest that in future research the concept of *chronogeneity* should be introduced (Georgiades et al., 2017). Chronogeneity is the heterogeneity of profiles of abilities in relation to the dimension of time. Children with autism are rarely completely stable over time. This means that one interesting way to look at the results we presented here should be to investigate the same profiles in a longitudinal perspective. It may be the case that children that are in one profile at a given time could actually change profiles. Moreover, the fact that children with ASD can change in severity and profiles of abilities could be the reflection of specific outcomes of therapies and care support. A longitudinal approach may help show the evolution of profiles across time and different therapies.

Turning to the clinical implications of this study, we would like to highlight two main points. First, LITMUS tasks have been proved to be good tools for disentangling language abilities in several populations, and the present study showed that they can also be used to individuate language impairment in children with ASD. The fact that both SR and NWR allow quantitative and qualitative analyses of specific structures that are known to be difficult for children showing language impairment can be used by SLPs as a first screening of structural language abilities in children with ASD. Moreover, especially for LITMUS-SR, the variety of constructions included in the task design could be useful for better distinguishing the difficulties and the errors across computationally complex features in groups with and without language impairment. In addition, LITMUS-tasks can be used to describe structural language abilities on an individual level as well. We are mainly thinking about the distinction across the ASD-LN groups, between children showing some selective shortcomings on structures requiring high processing of computational complexity (relative clauses). Finally, since it has already been shown that these tasks are useful endophenotypes for language impairment in other clinical populations (e.g. SLI) and in bilingual children, they could be particularly useful in clinical practice for comparing profiles of language abilities across patients. We would recommend use of LITMUS tasks in clinical practice, suggesting that Identical repetition and Target structure scores be used as a first initial screening and in-depth error analysis as a way of individuating possible patterns of impairment.

Second, our study has highlighted the importance of considering the existence of multiple profiles of structural language/NV abilities in ASD, which do not display the same structural language/NV characteristics and consequently cannot be treated homogeneously. Heterogeneity is still a hallmark of ASD but with this analysis and a methodological selection of measures of language and cognition, we may have made a step forward into defining possible patterns of abilities. This could help both clinicians and researchers look deeply into which are the real profiles of language abilities following proper grouping. In conclusion, subtyping would increase the possibility of developing treatments tailored to the specific needs of individuals, based on their particular patterns of strengths and impairment.

Conclusions

A diagnosis of ASD includes specification of any co-occurrence with language impairment and/or cognitive disabilities. However, few studies have explicitly explored possible combinations of language (dis)ability and cognitive (dis)ability, with the aim of better defining profiles of structural language and nonverbal cognitive abilities. Our study proposed a systematic investigation of both of these abilities in a group of fifty-one 6- to 12-year-old children with ASD, monolingual and bilingual, based on explicitly motivated measures. Tasks of Sentence Repetition (SR) and Nonword Repetition (NWR) were argued to be the most likely to appropriately assess structural language abilities in children with ASD, while Raven's Progressive Matrices, Block Design and Matrix Reasoning subtests of the WISC-IV were chosen as the most appropriate nonverbal (NV) tasks to assess cognitive abilities in individuals with autism in relation with linguistic abilities. The use of a comprehensive, empirical model based on unsupervised machine learning, cluster analysis, allowed the identification of five profiles of structural language / NV abilities for morphosyntactic and phonological skills in both monolingual and bilingual children with ASD. Among these five profiles, all four logically possible combinations of structural language and NV abilities were detected. Crucially the existence of a double dissociation, , the ASD-LI with low NVIQ profile and , ASD-LN with low NVIQ profile support the idea of a separate language module in the brain. Moreover, the existence of individuals who can develop normal langauge abilities in the presence of impaired NV capacities contradicts any position that insists on cognitive prerequisites for positive langauge development. The comparison of the phenotypical linguistic realisation of the five, with a group of age-matched children with SLI (*n* = 26) and a group of TD children (4- to 12- year-olds, *n* = 84), questioned the legitimacy of the hypothesis that there is an SLI profile in ASD and the assumption that children with ASD-LN display spared structural language abilities. Our results showed that the language impaired children with ASD were more severely impaired than their SLI peers on morphosyntactic abilities, while some children with normal language abilities did not perform in line with their TD age-peers on complex syntactic structures. Finally, both the SR and the NWR tasks proved to be good indicators of language impairment in monolingual and bilingual children with ASD, highlighting the fact that bilingualism did not differently affect children with exposure to two languages.

We believe that our findings, together, illustrate rather clearly that progress in understanding language profiles in ASD is dependent on both wide investigation of the whole spectrum and use of robust structural language and nonverbal cognitive measures. Clinically speaking, subtyping would increase the possibility of developing therapies and care support tailored to the specific needs of individuals, based on their particular pattern of strengths and impairment.

Bibliography

Akshoomoff, N. (2006). Use of the Mullen Scales of Early Learning for the assessment of young children with autism spectrum disorders. *Child Neuropsychology*, *12*(4–5), 269–277.

Allen, D. A., & Rapin, I. (1980). Language disorders in preschool children: Predictors of outcome—A preliminary report—. *Brain and Development*, *2*(1), 73–80.

Altman, C., Armon-Lotem, S., Fichman, S., & Walters, J. (2016). Macrostructure, microstructure, and mental state terms in the narratives of English–Hebrew bilingual preschool children with and without specific language impairment. *Applied Psycholinguistics*, *37*(01), 165–193. https://doi.org/10.1017/S0142716415000466

Anderson, D. K., Lord, C., Risi, S., DiLavore, P. S., Shulman, C., Thurm, A., Pickles, A. (2007). Patterns of growth in verbal abilities among children with autism spectrum disorder. *Journal of Consulting and Clinical Psychology*, *75*(4), 594–604. https://doi.org/10.1037/0022-006X.75.4.594

Anthony, A., Bogle, D., Ingram, T. T. S., & McIsaac, M. (1971). *The Edinburgh Articulation Test*. Edinburgh and London.

APA. (2013). *Diagnostic and statistical manual of mental disorders (5th ed.)*. Washington DC.

Archibald, L., & Gathercole, S. (2007). The complexities of complex memory span: Storage and processing deficits in specific language impairment☆. *Journal of Memory and Language*, *57*(2), 177–194. https://doi.org/10.1016/j.jml.2006.11.004

Armon-Lotem, S., de Jong, J., & Meir, N. (2015). *Assessing multilingual children: Disentangling bilingualism from language impairment*.

Armon-Lotem, Sharon, & Meir, N. (2016). Diagnostic accuracy of repetition tasks for the identification of specific language impairment (SLI) in bilingual children: evidence from Russian and Hebrew: Accuracy of repetition tasks for identification of SLI in bilingual children. *International Journal of Language & Communication Disorders*, *51*(6), 715–731. https://doi.org/10.1111/1460-6984.12242

Arnold, J. E., Bennetto, L., & Diehl, J. J. (2009). Reference production in young speakers with and without autism: Effects of discourse status and processing constraints. *Cognition*, *110*(2), 131–146. https://doi.org/10.1016/j.cognition.2008.10.016

ASHA. (2004). *Preferred Practice Patterns for the Profession of Speech-Language Pathology* (No. PP2004- 00191). Rockville, MD: American Speech-Language-Hearing Association. https://doi.org/10.1044/policy.PP2004-00191

Asperger, H. (1944). Die „Autistischen Psychopathen” im Kindesalter. *European Archives of Psychiatry and Clinical Neuroscience*, *117*(1), 76–136.

Baayen, R. H. (2008). *Analyzing linguistic data. A practical introduction to statistics*. Cambridge: Cambridge University Press.

Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences*, *4*(11), 417–423.

Baddeley, A. D., Gathercole, S., & Watson, S. (2004). *The Nonword Memory Test*. Univeristy of Durham: UK.

Baio, J., Wiggins, L., Christensen, D. L., Maenner, M. J., Daniels, J., Warren, Z., Dowling, N. F. (2018). Prevalence of Autism Spectrum Disorder Among Children Aged 8 Years — Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2014. *MMWR. Surveillance Summaries*, *67*(6), 1–23. https://doi.org/10.15585/mmwr.ss6706a1

Baldimtsi, E., Peristeri, E., Tsimpli, I. M., & Nicolopoulou, A. (2016). Bilingual Children with High Functioning Autism Spectrum Disorder: Evidence from Oral Narratives and Non-verbal Executive Function Tasks. In J. Scott & D. Waughtal (Eds.) *Proceedings of the 40th annual Boston University Conference on Language Development*. Somerville, MA: Cascadilla (pp. 18–31).

Baltaxe, C. A. (1977). Pragmatic deficits in the language of autistic adolescents. *Journal of Pediatric Psychology,* pp. 176–180.

Baltaxe, C. A., & D'Angiola, N. (1996). Referencing skills in children with autism and specific language impairment. *International Journal of Language & Communication Disorders*, *31*(3), 245–258.

Barbeau, E. B., Soulières, I., Dawson, M., Zeffiro, T. A., & Mottron, L. (2013). The level and nature of autistic intelligence III: Inspection time. *Journal of Abnormal Psychology*, *122*(1), 295–301. https://doi.org/10.1037/a0029984

Barger, B. D., Campbell, J. M., & McDonough, J. D. (2013). Prevalence and Onset of Regression within Autism Spectrum Disorders: A Meta-analytic Review. *Journal of Autism and Developmental Disorders*, *43*(4), 817–828. https://doi.org/10.1007/s10803-012-1621-x

Baron-Cohen, S. (1988). Social and pragmatic deficits in autism: cognitive or affective? *Journal of Autism and Developmental Disorders*, *18*(3), 379–402.

Baron-Cohen, S., & Staunton, R. (1994). Do children with autism acquire the phonology of their peers? An examination of group identification through the window of bilingualism. *First Language*, *14*(42–43), 241–248. https://doi.org/10.1177/014272379401404216

Bartak, L., Rutter, M., & Cox, A. (1975). A comparative study of infantile autism and specific developmental receptive language disorder. *The British Journal of Psychiatry*, *126*(2), 127–145.

Bartolucci, G., & Pierce, S. J. (1977). A preliminary comparison of phonological development in autistic, normal, and mentally retarded subjects. *International Journal of Language & Communication Disorders*, *12*(2), 137–147.

Baxter, A. J., Brugha, T. S., Erskine, H. E., Scheurer, R. W., Vos, T., & Scott, J. G. (2015). The epidemiology and global burden of autism spectrum disorders. *Psychological Medicine*, *45*(03), 601–613. https://doi.org/10.1017/S003329171400172X

Bennett, T. A., Szatmari, P., Georgiades, K., Hanna, S., Janus, M., Georgiades, S. Thompson, A. (2014). Language Impairment and Early Social Competence in Preschoolers with Autism Spectrum Disorders: A Comparison of DSM-5 Profiles. *Journal of Autism and Developmental Disorders*, *44*(11), 2797–2808. https://doi.org/10.1007/s10803-014-2138-2

Bennett, T., Szatmari, P., Bryson, S., Volden, J., Zwaigenbaum, L., Vaccarella, Boyle, M. (2008). Differentiating Autism and Asperger Syndrome on the Basis of Language Delay or Impairment. *Journal of Autism and Developmental Disorders*, *38*(4), 616–625. https://doi.org/10.1007/s10803-007-0428-7

Bialystok, E. (2001). *Bilingualism in development: Language, literacy, and cognition* Cambridge University Press.

Bialystok, Ellen, & Martin, M. M. (2004). Attention and inhibition in bilingual children: evidence from the dimensional change card sort task. *Developmental Science*, *7*(3), 325–339. https://doi.org/10.1111/j.1467-7687.2004.00351.x

Bishop, D. V. M. (1998). Development of the Children's Communication Checklist (CCC): A method for assessing qualitative aspects of communicative impairment in children. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, *39*(6), 879–891.

Bishop, D. V. M. (2001). Genetic and environmental risks for specific language impairment in children. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *356*, 369–380.

Bishop, D. V. M. (2003a). Autism and specific language impairment: categorical distinction or continuum. *Autism: Neural Basis and Treatment Possibilities*, 213–234.

Bishop, D. V. M. (2003b). *TROG-2 test for reception of grammar-2*. London: Harcourt.

Bishop, D. V. M. (2006). What Causes Specific Language Impairment in Children? *Current Directions in Psychological Science*, *15*(5), 217–221.

Bishop, D. V. M. (2010). Overlaps Between Autism and Language Impairment: Phenomimicry or Shared Etiology? *Behavior Genetics*, *40*(5), 618–629. https://doi.org/10.1007/s10519-010-9381-x

Bishop, D. V. M., Snowling, M. J., Thompson, P. A., Greenhalgh, T., & and the CATALISE-2 consortium. (2017). Phase 2 of CATALISE: a multinational and multidisciplinary Delphi consensus study of problems with language development: Terminology. *Journal of Child Psychology and Psychiatry*, *58*(10), 1068–1080. https://doi.org/10.1111/jcpp.12721

Bishop, Dorothy V.M., Maybery, M., Wong, D., Maley, A., Hill, W., & Hallmayer, J. (2004). Are phonological processing deficits part of the broad autism phenotype? *American Journal of Medical Genetics*, *128B*(1), 54–60. https://doi.org/10.1002/ajmg.b.30039

Bishop, S. L., Guthrie, W., Coffing, M., & Lord, C. (2011). Convergent Validity of the Mullen Scales of Early Learning and the Differential Ability Scales in Children With Autism Spectrum Disorders. *American Journal on Intellectual and Developmental Disabilities*, *116*(5), 331–343. https://doi.org/10.1352/1944-7558-116.5.331

Blom, E., Boerma, T., Bosma, E., Cornips, L., & Everaert, E. (2017). Cognitive Advantages of Bilingual Children in Different Sociolinguistic Contexts. *Frontiers in Psychology*, (*8)* 1-12. https://doi.org/10.3389/fpsyg.2017.00552

Bölte, S., Dziobek, I., & Poustka, F. (2009). Brief Report: The Level and Nature of Autistic Intelligence Revisited. *Journal of Autism and Developmental Disorders*, *39*(4), 678–682. https://doi.org/10.1007/s10803-008-0667-2

Bortolini, U., Arfé, B., Caselli, C. M., Degasperi, L., Deevy, P., & Leonard, L. B. (2006). Clinical markers for specific language impairment in Italian: the contribution of clitics and non-word repetition. *International Journal of Language & Communication Disorders*, *41*(6), 695–712. https://doi.org/10.1080/13682820600570831

Botting, N., & Conti-Ramsden, G. (2003). Autism, primary pragmatic difficulties, and specific language impairment: Can we distinguish them using psycholinguistic markers? *Developmental Medicine & Child Neurology*, *45*(8), 515–524.

Boucher, J. (1976). Articulation in early childhood autism. *Journal of Autism and Childhood Schizophrenia*, *6*(4), 297–302.

Boucher, J. (2003). Language development in autism. *International Journal of Pediatric Otorhinolaryngology*, *67*, S159–S163. https://doi.org/10.1016/j.ijporl.2003.08.016

Brock, G., Pihur, V., Datta, S., & Datta, S. (2011). clValid, an R package for cluster validation. *Journal of Statistical Software (Brock et Al., March 2008)*.

Brynskov, C., Eigsti, I.-M., Jørgensen, M., Lemcke, S., Bohn, O.-S., & Krøjgaard, P. (2017). Syntax and Morphology in Danish-Speaking Children with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, *47*(2), 373–383. https://doi.org/10.1007/s10803-016-2962-7

Buuren, S. van, & Groothuis-Oudshoorn, K. (2010). mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software*, 1–68.

Camaioni, L., Perucchini, P., Muratori, F., Parrini, B., & Cesari, A. (2003). The communicative use of pointing in autism: developmental profile and factors related to change. *European Psychiatry*, *18*(1), 6–12. https://doi.org/10.1016/S0924-9338(02)00013-5

Cantiani, C., Choudhury, N. A., Yu, Y. H., Shafer, V. L., Schwartz, R. G., & Benasich, A. A. (2016). From Sensory Perception to Lexical-Semantic Processing: An ERP Study in Non-Verbal Children with Autism. *PLOS ONE*, *11*(8), e0161637. https://doi.org/10.1371/journal.pone.0161637

Capps, L., Kehers, J., & Sigman, M. (1998). Conversational abilities among children with autism and children with developmental delays. *Autism*, pp. 325–344.

CDC. (2014). Prevalence of autism spectrum disorder among children aged 8 years-autism and developmental disabilities monitoring network, 11 sites, United States, 2010. *Morbidity and Mortality Weekly Report. Surveillance Summaries.* Washington, DC, *63*(2), 1, pp 1-21.

Charman, T., Pickles, A., Simonoff, E., Chandler, S., Loucas, T., & Baird, G. (2011). IQ in children with autism spectrum disorders: data from the Special Needs and Autism Project (SNAP). *Psychological Medicine*, *41*(03), 619–627. https://doi.org/10.1017/S0033291710000991

Charman, Tony, Baron-Cohen, S., Swettenham, J., Baird, G., Drew, A., & Cox, A. (2003). Predicting language outcome in infants with autism and pervasive developmental disorder. *International Journal of Language & Communication Disorders*, *38*(3), 265–285. https://doi.org/10.1080/136820310000104830

Charman, Tony, Taylor, E., Drew, A., Cockerill, H., Brown, J.-A., & Baird, G. (2005). Outcome at 7 years of children diagnosed with autism at age 2: predictive validity of assessments conducted at 2 and 3 years of age and pattern of symptom change over time. *Journal of Child Psychology and Psychiatry*, *46*(5), 500–513. https://doi.org/10.1111/j.1469-7610.2004.00377.x

Cheung, H., & Kemper, S. (1993). Recall and articulation of English and Chinese words by Chinese-English bilinguals. *Memory & Cognition*, *21*(5), 666–670. https://doi.org/10.3758/BF03197197

Chiat, S. (2015). Nonword repetition. In S. Armon-Lotem, J. de Jong & N. Meir (Eds.), *Methods for Assessing Multilingual Children: Disentangling Bilingualism from Language Impairment*. Bristol: UK, pp. 125–150.

Chiat, S., Armon-Lotem, S., Marinis, T., Polišenská, K., Roy, P., & Seeff-Gabriel, B. (2013). The potential of sentence imitation tasks for assessment of language abilities in sequential bilingual children. In V. Mueller-Gathercole (Ed.) *Issues in the Assessment of Bilinguals*, Vol. Multilingual Matters, Bristol: UK, pp. 56–89.

Chin, C. E., Ledesma, H. M. L., Cirino, P. T., Sevcik, R. A., Morris, R. D., Frijters, J. C., & Lovett, M. W. (2001). Relation Between Kaufman Brief Intelligence Test and WISC-III Scores of Children with RD. *Journal of Learning Disabilities*, *34*(1), 2–8. https://doi.org/10.1177/002221940103400101

Chita-Tegmark, M., Arunachalam, S., Nelson, C. A., & Tager-Flusberg, H. (2015). Eye-Tracking Measurements of Language Processing: Developmental Differences in Children at High Risk for ASD. *Journal of Autism and Developmental Disorders*, *45*(10), 3327–3338. https://doi.org/10.1007/s10803-015-2495-5

Chloe R. Marshall, & Heather K. J. van der Lely. (2009). Effects of Word Position and Stress on Onset Cluster Production: Evidence from Typical Development, Specific Language Impairment, and Dyslexia. *Language*, *85*(1), 39–57. https://doi.org/10.1353/lan.0.0081

Chomsky, N. (1965). *Aspacts of the Theory of Syntax*. MIT Cambridge Research Lab of Electronics.

Chomsky, N. (1980). *Rules and Representations*. New York: Columbia University Press.

Cleland, J., Gibbon, F. E., Peppé, S. J. E., O'Hare, A., & Rutherford, M. (2010). Phonetic and phonological errors in children with high functioning autism and Asperger syndrome. *International Journal of Speech-Language Pathology*, *12*(1), 69–76. https://doi.org/10.3109/17549500903469980

Coady, J. A., & Evans, J. L. (2008). Uses and interpretations of non-word repetition tasks in children with and without specific language impairments (SLI). *International Journal of Language & Communication Disorders*, *43*(1), 1–40. https://doi.org/10.1080/13682820601116485

Contemori, C., & Belletti, A. (2014). Relatives and passive object relatives in Italian-speaking children and adults: Intervention in production and comprehension. *Applied Psycholinguistics*, *35*(06), 1021–1053. https://doi.org/10.1017/S0142716412000689

Conti-Ramsden, G., Botting, N., & Faragher, B. (2001). Psycholinguistic Markers for Speci®c Language Impairment (SLI). *Journal of Child Psychology and Psychiatry*, *42*(6), 741–748.

Conti-Ramsden, G., Simkin, Z., & Botting, N. (2006). The prevalence of autistic spectrum disorders in adolescents with a history of specific language impairment (SLI). *Journal of Child Psychology and Psychiatry*, *47*(6), 621–628. https://doi.org/10.1111/j.1469-7610.2005.01584.x

Courchesne, V., Meilleur, A.-A. S., Poulin-Lord, M.-P., Dawson, M., & Soulières, I. (2015). Autistic children at risk of being underestimated: school-based pilot study of a strength-informed assessment. *Molecular Autism*, *6*(1), 6-12. https://doi.org/10.1186/s13229-015-0006-3

Courchesne, V., Nader, A.-M., Girard, D., Bouchard, V., Danis, É., & Soulières, I. (2016). Le profil cognitif au service des apprentissages : optimiser le potentiel des enfants sur le spectre de l'autisme. *Revue québécoise de psychologie*, *37*(2), 141. https://doi.org/10.7202/1040041ar

Crawford, J. R., Garthwaite, P. H., & Porter, S. (2010). Point and interval estimates of effect sizes for the case-controls design in neuropsychology: Rationale, methods, implementations, and proposed reporting standards. *Cognitive Neuropsychology*, (27), 245–260.

Dawson, M., Soulières, I., Ann Gernsbacher, M., & Mottron, L. (2007). The level and nature of autistic intelligence. *Psychological Science*, *18*(8), 657–662.

de Almeida, L., Ferré, S., Morin, E., Prévost, P., dos Santos, C., Tuller, L., … Barthez, M.-A. (2017). Identification of bilingual children with Specific Language Impairment in France. *Linguistic Approaches to Bilingualism*, *7*(3–4), 331–358. https://doi.org/10.1075/lab.15019.alm

de Bildt, A., Oosterling, I. J., van Lang, N. D. J., Sytema, S., Minderaa, R. B., van Engeland, H., … de Jonge, M. V. (2011). Standardized ADOS Scores: Measuring Severity of Autism Spectrum Disorders in a Dutch Sample. *Journal of Autism and Developmental Disorders*, *41*(3), 311–319. https://doi.org/10.1007/s10803-010-1057-0

De Villiers, J., Myers, B., & Stainton, R. J. (2010). Differential pragmatic abilities and autism spectrum disorders: The case of pragmatic determinants of literal content. *Pragmatics in Context. Toronto: Antares*. Retrieved from https://works.bepress.com/robertstainton/104/download/

Delage, H., Monjauze, C., Hamann, C., & Tuller, L. (2008). Relative Clauses in Atypical Acquisition of French. In A. Gavarró, M. João Freitas (Eds.) *Language acquisition and development: Proceedings of GALA*, pp. 166–176.

Devescovi, A., & Cristina Caselli, M. (2007). Sentence repetition as a measure of early grammatical development in Italian. *International Journal of Language & Communication Disorders*, *42*(2), 187–208. https://doi.org/10.1080/13682820601030686

Dewey, M. A., & Everard, M. P. (1974). The near-normal autistic adolescent. *Journal of Autism and Developmental Disorders*, *4*(4), 348–356.

Diehl, Joshua J., Bennetto, L., Watson, D., Gunlogson, C., & McDonough, J. (2008). Resolving ambiguity: A psycholinguistic approach to understanding prosody processing in high-functioning autism. *Brain and Language*, *106*(2), 144–152. https://doi.org/10.1016/j.bandl.2008.04.002

Diehl, Joshua John, Friedberg, C., Paul, R., & Snedeker, J. (2015). The use of prosody during syntactic processing in children and adolescents with autism spectrum disorders. *Development and Psychopathology*, *27*(03), 867–884. https://doi.org/10.1017/S0954579414000741

Dispaldro, M., Leonard, L. B., & Deevy, P. (2013). Real-Word and Nonword Repetition in Italian-Speaking Children With Specific Language Impairment: A Study of Diagnostic Accuracy. *Journal of Speech Language and Hearing Research*, *56*(1), 323-336. https://doi.org/10.1044/1092-4388(2012/11-0304)

Dollaghan, C., & Campbell, T. F. (1998). Nonword Repetition and Child Language Impairment. *Journal of Speech Language and Hearing Research*, *41*(5), 1136. https://doi.org/10.1044/jslhr.4105.1136

dos Santos, C., & Ferré, S. (2018). A Nonword Repetition Task to Assess Bilingual Children’s Phonology. *Language Acquisition*, *25*(1), 58–71. https://doi.org/10.1080/10489223.2016.1243692

Dumont, R., Cruse, C. L., Price, L., & Whelley, P. (1996). The relationship between the Differential Ability Scales (DAS) and the Wechsler Intelligence Scale for Children? Third Edition (WISC-III) for students with learning disabilities. *Psychology in the Schools*, *33*(3), 203–220.

Dunn, M., & Dunn, L. M. (1997). *Peabody Picture Vocabulary Test-3*. Circle Pines, MN: AGS.

Eigsti, I.-M., & Bennetto, L. (2009). Grammaticality judgments in autism: Deviance or delay. *Journal of Child Language*, *36*(05), 999. https://doi.org/10.1017/S0305000909009362

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. https://doi.org/10.1007/s10803-006-0239-2

Eigsti, I.-M., de Marchena, A. B., Schuh, J. M., & Kelley, E. (2011). Language acquisition in autism spectrum disorders: A developmental review. *Research in Autism Spectrum Disorders*, *5*(2), 681–691. https://doi.org/10.1016/j.rasd.2010.09.001

Ellis Weismer, S., Davidson, M. M., Gangopadhyay, I., Sindberg, H., Roebuck, H., & Kaushanskaya, M. (2017). The role of nonverbal working memory in morphosyntactic processing by children with specific language impairment and autism spectrum disorders. *Journal of Neurodevelopmental Disorders*, *9*(1). https://doi.org/10.1186/s11689-017-9209-6

Ellis Weismer, S., Gernsbacher, M. A., Stronach, S., Karasinski, C., Eernisse, E. R., Venker, C. E., & Sindberg, H. (2011). Lexical and Grammatical Skills in Toddlers on the Autism Spectrum Compared to Late Talking Toddlers. *Journal of Autism and Developmental Disorders*, *41*(8), 1065–1075. https://doi.org/10.1007/s10803-010-1134-4

Evans, K. E., & Demuth, K. (2012). Individual differences in pronoun reversal: Evidence from two longitudinal case studies. *Journal of Child Language*, *39*(01), 162–191. https://doi.org/10.1017/S0305000911000043

Feltmate, K., & Rehab, C. (2008). Language Learning in Four Bilingual Children with Down Syndrome: A Detailed Analysis of Vocabulary and Morphosyntax, *Canadian Journal of Speech-Language Pathology and Audiology 32*(1), 6-20.

Fenson, L., Pethick, S., Renda, C., Cox, J. L., Dale, P. S., & Reznick, J. S. (2000). Short-form versions of the MacArthur communicative development inventories. *Applied Psycholinguistics*, *21*(1), 95–116.

Ferré, S., & dos Santos, C. (2015). *The LITMUS-NWR-French*. Université de Tours.

Ferré, S., Tuller, L., Sizaret, E., & Barthez, M.-A. (2012). Acquiring and avoiding phonological complexity in SLI vs. typical development of French: The case of consonant clusters. In P. Hoole, L. Bombien, M. Pouplier, C. Mooshammer, & B. Kühnert (Eds.), *Consonant Clusters and Structural Complexity*. Berlin, Boston: DE GRUYTER. https://doi.org/10.1515/9781614510772.285

Fleckstein, A., Prévost, P., Tuller, L., Sizaret, E., & Zebib, R. (2016). How to identify SLI in bilingual children: A study on sentence repetition in French. *Language Acquisition*, 1–17.

Fodor, J. A. (1985). Precis of the modularity of mind. *Behavioral and Brain Sciences*, *8*(1), 1–5.

Fox, M. C., & Charness, N. (2010). How to Gain Eleven IQ Points in Ten Minutes: Thinking Aloud Improves Raven's Matrices Performance in Older Adults. *Aging, Neuropsychology, and Cognition*, *17*(2), 191–204. https://doi.org/10.1080/13825580903042668

Franck, J., Cronel-Ohayon, S., Chillier, L., Frauenfelder, U. H., Hamann, C., Rizzi, L., & Zesiger, P. (2004). Normal and pathological development of subject–verb agreement in speech production: a study on French children. *Journal of Neurolinguistics*, *17*(2–3), 147–180. https://doi.org/10.1016/S0911-6044(03)00057-5

Friedmann, N., Belletti, A., & Rizzi, L. (2009). Relativized relatives: Types of intervention in the acquisition of A-bar dependencies. *Lingua*, *119*(1), 67–88. https://doi.org/10.1016/j.lingua.2008.09.002

Frisch, S. A., Large, N. R., & Pisoni, D. B. (2000). Perception of Wordlikeness: Effects of Segment Probability and Length on the Processing of Nonwords. *Journal of Memory and Language*, *42*(4), 481–496. https://doi.org/10.1006/jmla.1999.2692

Gathercole, S. E., Willis, C. S., Baddeley, A. D., & Emslie, H. (1994). The children's test of nonword repetition: A test of phonological working memory. *Memory*, *2*(2), 103–127. https://doi.org/10.1080/09658219408258940

Gathercole, V. C. M. (2002). Monolingual and bilingual acquisition: Learning different treatments of that-trace phenomena in English and Spanish. In D. Kimbrough Oller & R. E. Eilers (Eds.) *Language and literacy in bilingual children*. Clevedon: Multilingual Matters, pp 220-254.

Gathercole, V. C. M., Thomas, E. M., Roberts, E. J., Hughes, C. O., & Hughes, E. K. (2013). Why Assessment Needs to Take Exposure into Account: Vocabulary and Grammatical Abilities in Bilingual Children. In V. C. M. Gathercole (Ed.), *Issues in the Assessment of Bilinguals* (pp. 20–55). Bristol, Blue Ridge Summit: Multilingual Matters. https://doi.org/10.21832/9781783090105-004

Geneese, F., Paradis, J., & Crago, M. (2004). *Dual language development and disorders: A handbook on bilingualism and second language learning*. Baltimore, MD: Paul H. Brookes.

Georgiades, S., Bishop, S. L., & Frazier, T. (2017). Editorial Perspective: Longitudinal research in autism - introducing the concept of 'chronogeneity.' *Journal of Child Psychology and Psychiatry*, *58*(5), 634–636. https://doi.org/10.1111/jcpp.12690

Georgiades, S., Szatmari, P., Boyle, M., Hanna, S., Duku, E., Zwaigenbaum, Pathways in ASD Study Team. (2013). Investigating phenotypic heterogeneity in children with autism spectrum disorder: a factor mixture modeling approach: ASD factor mixture model. *Journal of Child Psychology and Psychiatry*, *54*(2), 206–215. https://doi.org/10.1111/j.1469-7610.2012.02588.x

Geurts, H. M., & Embrechts, M. (2008). Language Profiles in ASD, SLI, and ADHD. *Journal of Autism and Developmental Disorders*, *38*(10), 1931–1943. https://doi.org/10.1007/s10803-008-0587-1

Goad, H., & Rose, Y. (2004). Input elaboration, head faithfulness, and evidence for representation in the acquisition of left-edge clusters in West Germanic. In R. Kager, J. Pater, & W. Zonneveld (Eds.), *Constraints in Phonological Acquisition* (pp. 109–157). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511486418.005

Goldman, R., & Fristoe, M. (1986). *The Goldman-Fristoe Test of Articulation*. Circle Pines, MN: American Guidance Service.

Goldman, R., & Fristoe, M. (2000). *The Goldman-Fristoe Test of Articulation. Second Edition*. Circle Pines, MN: American Guidance Service.

Gonzalez-Barrero, A. M., & Nadig, A. (2017). Verbal fluency in bilingual children with Autism Spectrum Disorders. *Linguistic Approaches to Bilingualism*, *7*(3–4), 460–475. https://doi.org/10.1075/lab.15023.gon

Goodwin, A., Fein, D., & Naigles, L. R. (2012). Comprehension of Wh-Questions Precedes Their Production in Typical Development and Autism Spectrum Disorders: Wh-question comprehension precedes production. *Autism Research*, *5*(2), 109–123. https://doi.org/10.1002/aur.1220

Gotham, K., Pickles, A., & Lord, C. (2012). Trajectories of Autism Severity in Children Using Standardized ADOS Scores. *Pediatrics*, *130*(5), e1278–e1284. https://doi.org/10.1542/peds.2011-3668

Gotham, Katherine, Pickles, A., & Lord, C. (2009). Standardizing ADOS Scores for a Measure of Severity in Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*, *39*(5), 693–705. https://doi.org/10.1007/s10803-008-0674-3

Grimm, A., & Schulz, P. (2014). Specific Language Impairment and Early Second Language Acquisition: The Risk of Over- and Underdiagnosis. *Child Indicators Research*, *7*(4), 821–841. https://doi.org/10.1007/s12187-013-9230-6

Grundy, J. G., & Timmer, K. (2017). Bilingualism and working memory capacity: A comprehensive meta-analysis. *Second Language Research*, *33*(3), 325–340. https://doi.org/10.1177/0267658316678286

Guasti, M. T., & Cardinaletti, A. (2003). Relative clause formation in Romance childs production. *Probus*, *15*(1). https://doi.org/10.1515/prbs.2003.005

Guthrie, W., Swineford, L. B., Nottke, C., & Wetherby, A. M. (2013). Early diagnosis of autism spectrum disorder: stability and change in clinical diagnosis and symptom presentation: Stability of early diagnosis and symptoms in ASD. *Journal of Child Psychology and Psychiatry*, *54*(5), 582–590. https://doi.org/10.1111/jcpp.12008

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. https://doi.org/10.5565/rev/catjl.82

Hambly, C., & Fombonne, E. (2012). The Impact of Bilingual Environments on Language Development in Children with Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*, *42*(7), 1342–1352. https://doi.org/10.1007/s10803-011-1365-z

Happé, F. G. (1994). An advanced test of theory of mind: Understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *Journal of Autism and Developmental Disorders*, *24*(2), 129–154.

Happe, F. G. E. (1995). The Role of Age and Verbal Ability in the Theory of Mind Task Performance of Subjects with Autism. *Child Development*, *66*(3), 843. https://doi.org/10.2307/1131954

Harper-Hill, K., Copland, D., & Arnott, W. (2013). Do spoken nonword and sentence repetition tasks discriminate language impairment in children with an ASD? *Research in Autism Spectrum Disorders*, *7*(2), 265–275. https://doi.org/10.1016/j.rasd.2012.08.015

Hayashi, M., Kato, M., Igarashi, K., & Kashima, H. (2008). Superior fluid intelligence in children with Asperger's disorder. *Brain and Cognition*, *66*(3), 306–310. https://doi.org/10.1016/j.bandc.2007.09.008

Hayes, B. (2004). Phonological acquisition in Optimality Theory: the early stages. *Constraints in phonological acquisition*, 158-203.

Hennig, C., & Liao, T. F. (2013). How to find an appropriate clustering for mixed-type variables with application to socio-economic stratification. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *62*(3), 309–369.

Hess, L. (2006). I would like to play but I don't know how: a case study of pretend play in autism. *Child Language Teaching and Therapy*, *22*(1), 97–116. https://doi.org/10.1191/0265659006ct299oa

Hilchey, M. D., & Klein, R. M. (2011). Are there bilingual advantages on nonlinguistic interference tasks? Implications for the plasticity of executive control processes. *Psychonomic Bulletin & Review*, *18*(4), 625–658. https://doi.org/10.3758/s13423-011-0116-7

Hill, A. P., van Santen, J., Gorman, K., Langhorst, B. H., & Fombonne, E. (2015). Memory in language-impaired children with and without autism. *Journal of Neurodevelopmental Disorders*, *7*(1). https://doi.org/10.1186/s11689-015-9111-z

Houskeeper, J. L. (2011). *WISC-IV Matrix Reasoning and Picture Concepts Subtests: Does the Use of Verbal Mediation Confound Measurement of Fluid Reasoning?* Pacific University.

Hudry, K., Leadbitter, K., Temple, K., Slonims, V., McConachie, H., Aldred, C., Charman, T. (2010). Preschoolers with autism show greater impairment in receptive compared with expressive language abilities. *International Journal of Language & Communication Disorders*, *45*(6), 681–690. https://doi.org/10.3109/13682820903461493

Hus, V., & Lord, C. (2013). Effects of Child Characteristics on the Autism Diagnostic Interview-Revised: Implications for Use of Scores as a Measure of ASD Severity. *Journal of Autism and Developmental Disorders*, *43*(2), 371–381. https://doi.org/10.1007/s10803-012-1576-y

Jakubovitz, C., & Tuller, L. (2008). Specific language impairment in French. In D. Ayoun (Ed.) *Language Learning & Language Teaching*, Vol. 21. Amsterdam: John Benjamins Publishing Company, pp. 97–133. https://doi.org/10.1075/lllt.21.08jak

Janke, V., & Perovic, A. (2015). Intact grammar in HFA? Evidence from control and binding. *Lingua*, *164*, 68–86. https://doi.org/10.1016/j.lingua.2015.06.009

Jarvis, L. H., Danks, J. H., & Merriman, W. E. (1995). The effect of bilingualism on cognitive ability: A test of the level of bilingualism hypothesis. *Applied Psycholinguistics*, *16*(03), 293. https://doi.org/10.1017/S0142716400007311

Jensen de López, K., Schroeder, K., & Gavarró, A. (2018). Successful passive sentence comprehension among Danish adolescents with autism spectrum disorders. *Autism & Developmental Language Impairments*, *3*, 239694151876123. https://doi.org/10.1177/2396941518761239

Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, *21*(1), 60–99.

Joseph, R. M., Tager-Flusberg, H., & Lord, C. (2002). Cognitive profiles and social-communicative functioning in children with autism spectrum disorder. *Journal of Child Psychology and Psychiatry*, *43*(6), 807–821.

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*. https://doi.org/10.3389/fpsyg.2017.00319

Kaland, N., Møller-Nielsen, A., Callesen, K., Mortensen, E. L., Gottlieb, D., & Smith, L. (2002). A new advanced test of theory of mind: evidence from children and adolescents with Asperger syndrome. *Journal of Child Psychology and Psychiatry*, *43*(4), 517–528.

Kalb, L. G., Law, J. K., Landa, R., & Law, P. A. (2010). Onset Patterns Prior to 36 Months in Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*, *40*(11), 1389–1402. https://doi.org/10.1007/s10803-010-0998-7

Kalnak, N., Peyrard-Janvid, M., Forssberg, H., & Sahlén, B. (2014). Nonword Repetition – A Clinical Marker for Specific Language Impairment in Swedish Associated with Parents' Language-Related Problems. *PLoS ONE*, *9*(2), e89544. https://doi.org/10.1371/journal.pone.0089544

Kanner, L (1946). Irrelevant and metaphorical language in early infantile autism. *American Journal of Psychiatry*, p. 242—246.

Kanner, L. (1973). The birth of early infantile autism. *Journal of Autism and Childhood Schizophrenia*, *3*(2), 93–95.

Kapalková, S., Polišenská, K., & Vicenová, Z. (2013). Non-word repetition performance in Slovak-speaking children with and without SLI: novel scoring methods: Non-word repetition: novel scoring methods. *International Journal of Language & Communication Disorders*, *48*(1), 78–89. https://doi.org/10.1111/j.1460-6984.2012.00189.x

Kassambara, A. (2017). *Practical guide to cluster analysis in R: Unsupervised machine learning*. STHDA, Vol. 1.

Kaufman, L., & Rousseeuw, P. (1990). Partitioning Around Medoids (Program PAM). In *Finding Groups in Data. An Introduction to Cluster Analysis*. Vol. 344). John Wiley & Sons, Hoboken: New Jersey, pp. 68–125.

Kay-Raining Bird, E., Lamond, E., & Holden, J. (2012). Survey of bilingualism in autism spectrum disorders: Bilingualism and autism: a survey study. *International Journal of Language & Communication Disorders*, *47*(1), 52–64. https://doi.org/10.1111/j.1460-6984.2011.00071.x

Kay-Raining Bird, E., Trudeau, N., & Sutton, A. (2016). Pulling it all together: The road to lasting bilingualism for children with developmental disabilities. *Journal of Communication Disorders*, *63*, 63–78. https://doi.org/10.1016/j.jcomdis.2016.07.005

Kelley, E., Paul, J., Fein, D., & Naigles, L. (2006). Residual language deficits in optimal outcome children with a history of autism. *Journal of Autism and Developmental Disorders*, *36*(6), 807–827.

Kenworthy, L., Wallace, G. L., Powell, K., Anselmo, C., Martin, A., & Black, D. O. (2012). Early language milestones predict later language, but not autism symptoms in higher functioning children with autism spectrum disorders. *Research in Autism Spectrum Disorders*, *6*(3), 1194–1202. https://doi.org/10.1016/j.rasd.2012.03.009

Khetrapal, N., & Thornton, R. (2017). C-Command in the Grammars of Children with High Functioning Autism. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.00402

Khomsi, A. (2001). ELO: évaluation du langage oral. ECPA, Les Editions du Centre de psychologie appliquée.

Khomsi, A., Khomsi, J., Pasquet, F., & Parbeau-Guéno, A. (2007). Bilan Informatisé de Langage Oral au cycle 3 et au Collège (BILO-3C)[Computerized assessment of oral Language for 8–10 years old children, and at middle school]. Paris: Editions 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. https://doi.org/10.1080/01690960042000058

Klingner, J. K., & Artiles, A. (2003). When should bilingual student be in special education? *Educational Leadership*, *61*(2), 66–71.

Kohl, M., Beauquier-Maccotta, B., Bourgeois, M., Clouard, C., Donde, S., Mosser, A., Robel, L. (2008). Bilinguisme et troubles du langage chez l'enfant : étude rétrospective. *La psychiatrie de l'enfant*, *51*(2), 577. https://doi.org/10.3917/psye.512.0577

Korkman, M., Kirk, U., & Kemp, S. (2007). *Nepsy-II*. San Antonio, TX: Pearson.

Kover, S. T., Edmunds, S. R., & Ellis Weismer, S. (2016). Brief Report: Ages of Language Milestones as Predictors of Developmental Trajectories in Young Children with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, *46*(7), 2501–2507. https://doi.org/10.1007/s10803-016-2756-y

Kremer-Sadlik, T. (2005). To Be or Not to Be Bilingual: Autistic Children from Multilingual Families. In J. Cohen, K. T. McAlister, K. Rolstad and J. MacSwan (Eds.) *ISB4: Proceedings of the 4th International Symposium on Bilingualism*. Sommerville, MA:Casadilla Press, pp. 1225–1234.

Landa, R. J., Holman, K. C., & Garrett-Mayer, E. (2007). Social and Communication Development in Toddlers With Early and Later Diagnosis of Autism Spectrum Disorders. *Archives of General Psychiatry*, *64*(7), 853. https://doi.org/10.1001/archpsyc.64.7.853

Le Couteur, A., Lord, C., & Rutter, M. (2003). *The Autism Diagnostic Interview-Revised (ADI-R)*. Los Angeles, CA: Western Psychological Services.

Leclercq, A.-L., Quémart, P., Magis, D., & Maillart, C. (2014). The sentence repetition task: A powerful diagnostic tool for French children with specific language impairment. *Research in Developmental Disabilities*, *35*(12), 3423–3430. https://doi.org/10.1016/j.ridd.2014.08.026

Léger, E. (2017). *Interprétation des pronoms clitiques objets chez les enfants avec TSA et chez les enfants avec TSL. Étude comparative en suivi du regard*. Université de Tours.

Leisch, F. (2006). A Toolbox for K-Centroids Cluster Analysis. *Computational Statistics & Data Analysis*, *51*(2), 526–544. https://doi.org/10.1016/j.csda.2005.10.006

Lelord, G., & Barthélémy, C. (2003). *Echelle d'evaluation des comportements autistiques, ECAR*. ECAN.

Leyfer, O. T., Tager-Flusberg, H., Dowd, M., Tomblin, J. B., & Folstein, S. E. (2008). Overlap between autism and specific language impairment: comparison of Autism Diagnostic Interview and Autism Diagnostic Observation Schedule scores. *Autism Research*, *1*(5), 284–296. https://doi.org/10.1002/aur.43

Lindgren, K. A., Folstein, S. E., Tomblin, J. B., & Tager-Flusberg, H. (2009). Language and reading abilities of children with autism spectrum disorders and specific language impairment and their first-degree relatives. *Autism Research*, *2*(1), 22–38. https://doi.org/10.1002/aur.63

Lippke, B. A., Dickey, S. E., Selmar, J. W., & Soder, A. L. (1997). *PAT-3: Photo Articulation Test.* London: Pearson, Pro-Ed.

Lloyd, H., Paintin, K., & Botting, N. (2006). Performance of children with different types of communication impairment on the Clinical Evaluation of Language Fundamentals (CELF). *Child Language Teaching and Therapy*, *22*(1), 47–67. https://doi.org/10.1191/0265659006ct297oa

Lombardi, & Potter. (1992). The Regeneration of Syntax in Short Term Memory. *Journal of Memory and Language*, *31*, 13-33.

Lord, C., Rutter, M., DiLavore, P. S., & Risi, S. (2003). *Autism Diagnostic Observation Schedule: ADOS*. Los Angeles, CA: Western Psychological Services.

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

Loucas, T., Charman, T., Pickles, A., Simonoff, E., Chandler, S., Meldrum, D., & Baird, G. (2008). Autistic symptomatology and language ability in autism spectrum disorder and specific language impairment. *Journal of Child Psychology and Psychiatry*, *49*(11), 1184–1192.

Luyster, R. J., Kadlec, M. B., Carter, A., & Tager-Flusberg, H. (2008). Language Assessment and Development in Toddlers with Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*, *38*(8), 1426–1438. https://doi.org/10.1007/s10803-007-0510-1

Lyall, K., Croen, L., Daniels, J., Fallin, M. D., Ladd-Acosta, C., Lee, B. K., (2017). The changing epidemiology of autism spectrum disorders. *Annual Review of Public Health*, *38*, 81–102.

MacKay, G., & Shaw, A. (2004). A comparative study of figurative language in children with autistic spectrum disorders. *Child Language Teaching and Therapy*, *20*(1), 13–32.

Mackintosh, N., & Mackintosh, K. J. (2011). *IQ and human intelligence*. Oxford University Press: Oxford.

Maddieson, I. (2009). Calculating phonological complexity. In F. Pellegrino (Ed.) *Approaches to Phonological Complexity*. Berlin: Walter de Gruyter. https://doi.org/10.1515/9783110223958

Månsson, J., Stjernqvist, K., Serenius, F., Ådén, U., & Källén, K. (2018). Agreement Between Bayley-III Measurements and WISC-IV Measurements in Typically Developing Children. *Journal of Psychoeducational Assessment*, 1-14. https://doi.org/10.1177/0734282918781431

Marinis, T., Armon-Lotem, S., & Pontikas, G. (2017). Language impairment in bilingual children: State of the art 2017. *Linguistic Approaches to Bilingualism*, *7*(3–4), 265–276. https://doi.org/10.1075/lab.00001.mar

Mayes, S.D., & Calhoun, S. L. (2003). Analysis of WISC-III, Stanford-Binet-IV, and Academic Achievement Test Scores in Children with Autism. *Journal of Autism and Developmental Disorders*, *33*(3), 329-341.

Mayes, Susan Dickerson, & Calhoun, S. L. (2008). WISC-IV and WIAT-II Profiles in Children With High-Functioning Autism. *Journal of Autism and Developmental Disorders*, *38*(3), 428–439. https://doi.org/10.1007/s10803-007-0410-4

Mayo, J., Chlebowski, C., Fein, D. A., & Eigsti, I.-M. (2013). Age of First Words Predicts Cognitive Ability and Adaptive Skills in Children with ASD. *Journal of Autism and Developmental Disorders*, *43*(2), 253–264. https://doi.org/10.1007/s10803-012-1558-0

McCleery, J. P., Tully, L., Slevc, L. R., & Schreibman, L. (2006). Consonant production patterns of young severely language-delayed children with autism. *Journal of*

Communication Disorders, *39*(3), 217–231. https://doi.org/10.1016/j.jcomdis.2005.12.002

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. https://doi.org/10.1007/s10803-011-1210-4

McPartland, J. C., Reichow, B., & Volkmar, F. R. (2012). Sensitivity and Specificity of Proposed DSM-5 Diagnostic Criteria for Autism Spectrum Disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, *51*(4), 368–383. https://doi.org/10.1016/j.jaac.2012.01.007

Mehling, M. H., & Tassé, M. J. (2016). Severity of Autism Spectrum Disorders: Current Conceptualization, and Transition to DSM-5. *Journal of Autism and Developmental Disorders*, *46*(6), 2000–2016. https://doi.org/10.1007/s10803-016-2731-7

Mervis, C. B., & Velleman, S. L. (2011). Children With Williams Syndrome: Language, Cognitive, and Behavioral Characteristics and Their Implications for Intervention. *Perspectives on Language Learning and Education*, *18*(3), 98. https://doi.org/10.1044/lle18.3.98

Minshew, N. J., Sweeney, J., & Luna, B. (2002). Autism as a selective disorder of complex information processing and underdevelopment of neocortical systems. *Molecular Psychiatry*, *7*(S2), S14.

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*. https://doi.org/10.3389/fpsyg.2017.00320

Moll, K., Hulme, C., Nag, S., & Snowling, M. J. (2015). Sentence repetition as a marker of language skills in children with dyslexia. *Applied Psycholinguistics*, *36*(02), 203–221. https://doi.org/10.1017/S0142716413000209

Morales, J., Calvo, A., & Bialystok, E. (2013). Working memory development in monolingual and bilingual children. *Journal of Experimental Child Psychology*, *114*(2), 187–202. https://doi.org/10.1016/j.jecp.2012.09.002

Morales-Hidalgo, P., Roigé-Castellví, J., Hernández-Martínez, C., Voltas, N., & Canals, J. (2018). Prevalence and Characteristics of Autism Spectrum Disorder Among Spanish School-Age Children. *Journal of Autism and Developmental Disorders*, *48*(9), 3176–3190. https://doi.org/10.1007/s10803-018-3581-2

Morton, B., & Carlson, S. (2017). The bilingual advantage: Evidence and alternative views. In M. Hoskyn, G. Iarocci, A. Young (Eds.) *Executive function in young children's everyday lives. A handbook for professionals in applied psychology*. New York: Oxford University Press online. http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780199980864.
001.0001/acprof-9780199980864-chapter-8.

Mostofsky, S. H., Goldberg, M. C., Landa, R. J., & Denckla, M. B. (2000). Evidence for a deficit in procedural learning in children and adolescents with autism: Implications for cerebellar contribution. *Journal of the International Neuropsychological Society*, *6*(7), 752–759. https://doi.org/10.1017/S1355617700677020

Mottron, L. (2004). Matching strategies in cognitive research with individuals with high-functioning autism: Current practices, instrument biases, and recommendations. *Journal of Autism and Developmental Disorders*, *34*(1), 19–27.

Mottron, L., Dawson, M., Soulières, I., Hubert, B., & Burack, J. (2006). Enhanced Perceptual Functioning in Autism: An Update, and Eight Principles of Autistic Perception. *Journal of Autism and Developmental Disorders*, *36*(1), 27–43. https://doi.org/10.1007/s10803-005-0040-7

Mullen, E. M. (1995). *Mullen scales of early learning*. Circle Pines, MN: AGS.

Munson, B., Kurtz, B. A., & Windsor, J. (2005). The Influence of Vocabulary Size, Phonotactic Probability, and Wordlikeness on Nonword Repetitions of Children With and Without Specific Language Impairment. *Journal of Speech Language and Hearing Research*, *48*(5), 1033. https://doi.org/10.1044/1092-4388(2005/072)

Nader, A.-M., Courchesne, V., Dawson, M., & Soulières, I. (2016). Does WISC-IV Underestimate the Intelligence of Autistic Children? *Journal of Autism and Developmental Disorders*, *46*(5), 1582–1589. https://doi.org/10.1007/s10803-014-2270-z

Nader, A.-M., Jelenic, P., & Soulières, I. (2015). Discrepancy between WISC-III and WISC-IV Cognitive Profile in Autism Spectrum: What Does It Reveal about Autistic Cognition? *PLOS ONE*, *10*(12), e0144645. https://doi.org/10.1371/journal.pone.0144645

Nadig, A., & Mulligan, A. (2017). Intact non-word repetition and similar error patterns in language-matched children with autism spectrum disorders: A pilot study. *Journal of Communication Disorders*, *66*, 13–21. https://doi.org/10.1016/j.jcomdis.2017.03.003

Naigles, L., & Fein, D. (2017). Looking through their eyes: Tracking early language comprehension in ASD. In L. R. Naigles (Ed.) *Language and the human lifespan series. Innovative investigations of language in autism spectrum disorder*. Washington, DC, US: American Psychological Association; Berlin, Germany: Walter de Gruyter GmbH, pp. 49–69.

Naigles, L. R., Cheng, M., Xu Rattanasone, N., 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. https://doi.org/10.1016/j.rasd.2016.03.008

Naigles, L. R., Kelty, E., Jaffery, R., & Fein, D. (2011). Abstractness and continuity in the syntactic development of young children with autism. *Autism Research*, *4*(6), 422–437. https://doi.org/10.1002/aur.223

Norbury, C. F. (2005). Barking up the wrong tree? Lexical ambiguity resolution in children with language impairments and autistic spectrum disorders. *Journal of Experimental Child Psychology*, *90*(2), 142–171. https://doi.org/10.1016/j.jecp.2004.11.003

Norbury, C. F., & Bishop, D. V. M. (2002). Inferential processing and story recall in children with communication problems: a comparison of specific language impairment, pragmatic language impairment and high-functioning autism. *International Journal of Language & Communication Disorders*, *37*(3), 227–251. https://doi.org/10.1080/13682820210136269

Norbury, C. F., Griffiths, H., & Nation, K. (2010). Sound before meaning: Word learning in autistic disorders. *Neuropsychologia*, *48*(14), 4012–4019. https://doi.org/10.1016/j.neuropsychologia.2010.10.015

Novogrodsky, R. (2013). Subject pronoun use by children with autism spectrum disorders (ASD). *Clinical Linguistics & Phonetics*, *27*(2), 85–93. https://doi.org/10.3109/02699206.2012.742567

Ohashi, J. K., Mirenda, P., Marinova-Todd, S., Hambly, C., Fombonne, E., Szatmari, P., Thompson, A. (2012). Comparing early language development in monolingual- and bilingual- exposed young children with autism spectrum disorders. *Research in Autism Spectrum Disorders*, *6*(2), 890–897. https://doi.org/10.1016/j.rasd.2011.12.002

Okanda, M., Moriguchi, Y., & Itakura, S. (2010). Language and Cognitive Shifting: Evidence from Young Monolingual and Bilingual Children. *Psychological Reports*, *107*(1), 68-78E. https://doi.org/10.2466/03.10.28.PR0.107.4.68-78

Oliveras-Rentas, R. E., Kenworthy, L., Roberson, R. B., Martin, A., & Wallace, G. L. (2012). WISC-IV Profile in High-Functioning Autism Spectrum Disorders: Impaired Processing Speed is Associated with Increased Autism Communication Symptoms and Decreased Adaptive Communication Abilities. *Journal of Autism and Developmental Disorders*, *42*(5), 655–664. https://doi.org/10.1007/s10803-011-1289-7

Ozonoff, S., Heung, K., Byrd, R., Hansen, R., & Hertz-Picciotto, I. (2008). The onset of autism: patterns of symptom emergence in the first years of life. *Autism Research*, *1*(6), 320–328. https://doi.org/10.1002/aur.53

Ozonoff, S., Strayer, D. L., McMahon, W. M., & Filloux, F. (1994). Executive function abilities in autism and Tourette syndrome: An information processing approach. *Journal of Child Psychology and Psychiatry*, *35*(6), 1015–1032.

Paradis, J. (2005). Grammatical Morphology in Children Learning English as a Second Language: Implications of Similarities With Specific Language Impairment, *Language, Speech and Hearing Sciences in Schools*(36), 172–187.

Paradis, J. (2007). Bilingual children with specific language impairment: Theoretical and applied issues. *Applied Psycholinguistics*, *28*(03). https://doi.org/10.1017/S0142716407070300

Paradis, J. (2010). The interface between bilingual development and specific language impairment. *Applied Psycholinguistics*, *31*(02), 227. https://doi.org/10.1017/S0142716409990373

Paradis, J. (2016). An agenda for knowledge-oriented research on bilingualism in children with developmental disorders. *Journal of Communication Disorders*, *63*, 79–84. https://doi.org/10.1016/j.jcomdis.2016.08.002

Paradis, J., & Crago, M. (2000). Tense and Temporality: A Comparison Between Children Learning a Second Language and Children With SLI. *Journal of Speech Language and Hearing Research*, *43*(4), 834. https://doi.org/10.1044/jslhr.4304.834

Park, C. J., Yelland, G. W., Taffe, J. R., & Gray, K. M. (2012a). Brief Report: The Relationship Between Language Skills, Adaptive Behavior, and Emotional and Behavior Problems in Pre-schoolers with Autism. *Journal of Autism and Developmental Disorders*, *42*(12), 2761–2766. https://doi.org/10.1007/s10803-012-1534-8

Park, C. J., Yelland, G. W., Taffe, J. R., & Gray, K. M. (2012b). Morphological and syntactic skills in language samples of pre school aged children with autism: Atypical development? *International Journal of Speech-Language Pathology*, *14*(2), 95–108. https://doi.org/10.3109/17549507.2011.645555

Park, S. (2014). Bilingualism and children with autism spectrum disorders: issues, research, and implications. *Nys Tesol Journal*, *1*(2), 122–129.

Peal, E., & Lambert, W. E. (1962). Peal, E., & Lambert, W. E. (1962). The relation of bilingualism to intelligence. Psychological Monographs: general and applied, 76(27), 1. *Psychological Monographs: General and Applied*, *76*(27(1)).

Pennington, B. F., & Ozonoff, S. (1996). Executive functions and developmental psychopathology. *Journal of Child Psychology and Psychiatry*, pp. 51–87.

Perner, J., & Wimmer, H. (1985). John thinks that Mary thinks that…" attribution of second-order beliefs by 5-to 10-year-old children. *Journal of Experimental Child Psychology*, pp. 437–471.

Perovic, A., Modyanova, N., & Wexler, K. (2013). Comparison of Grammar in Neurodevelopmental Disorders: The Case of Binding in Williams Syndrome and Autism With and Without Language Impairment. *Language Acquisition*, *20*(2), 133–154. https://doi.org/10.1080/10489223.2013.766742

Perron-Borelli, M. (1996). *Échelles différentielles d'efficience intellectuelle—forme révisée (EDEI-R)*. ECPA.

Petersen, J. M., Marinova-Todd, S. H., & Mirenda, P. (2012). Brief Report: An Exploratory Study of Lexical Skills in Bilingual Children with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, *42*(7), 1499–1503. https://doi.org/10.1007/s10803-011-1366-y

Polišenská, K., Chiat, S., & Roy, P. (2015). Sentence repetition: what does the task measure?: Sentence repetition: what does the task measure? *International Journal of Language & Communication Disorders*, *50*(1), 106–118. https://doi.org/10.1111/1460-6984.12126

Poncelet, M., & Van der Linden, M. (2003). L'évaluation du stock phonologique de la mémoire de travail: élaboration d'une épreuve de répétition de non-mots pour population francophone. *Revue de Neuropsychologie*, *13*(3), 377–407.

Potter, M. C., & Lombardi, L. (1990). Regeneration in the short-term recall of sentences. *Journal of Memory and Language*, *29*(6), 633–654. https://doi.org/10.1016/0749-596X(90)90042-X

Potter, M., & Lombardi, L. (1998). Syntactic priming in immediate racell of sentences. J*ournal of Memory and Language*, *38*, 265–282.

Prévost, P., Strik, N., & Tuller, L. (2014). *Wh*- questions in child L2 French: Derivational complexity and its interactions with L1 properties, length of exposure, age of exposure, and the input. *Second Language Research*, *30*(2), 225–250. https://doi.org/10.1177/0267658313519814

Prévost, P., Tuller, L., Barthez, M. A., Malvy, J., & Bonnet-Brilhault, F. (2017). Production and comprehension of French wh-questions by children with autism spectrum disorder: A comparative study with specific language impairment. *Applied Psycholinguistics*, *38*(05), 1095–1131. https://doi.org/10.1017/S0142716417000078

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*, 239694151879964. https://doi.org/10.1177/2396941518799643

Prevost, P., Zebib, R., & Tuller, L. (2012). *The LITMUS-SR-French*. Univerisité de Tours.

Prizant, B. M., & Duchan, J. F. (1981). The functions of immediate echolalia in autistic children. *Journal of Speech and Hearing Disorders*, *46*(3), 241–249.

Raiford, S. E., & Holdnack, J. A. (2014). *Technical and Interpretive Manual Supplement: Special Group Validity Studies With Other Measures and Additional Tables of WISC V*. London: Pearson, PsyhCorp.

Ramus, F. (2003). Theories of developmental dyslexia: insights from a multiple case study of dyslexic adults. *Brain*, *126*(4), 841–865. https://doi.org/10.1093/brain/awg076

Rapin, I., & Dunn, M. (1997). Language disorders in children with autism. *Seminars in Pediatric Neurology*, *4*(2), 86–92. https://doi.org/10.1016/S1071-9091(97)80024-1

Rapin, I., & Dunn, M. (2003). Update on the language disorders of individuals on the autistic spectrum. *Brain and Development*, *25*(3), 166–172. https://doi.org/10.1016/S0387-7604(02)00191-2

Rapin, I., Dunn, M. A., Allen, D. A., Stevens, M. C., & Fein, D. (2009). Subtypes of Language Disorders in School-Age Children With Autism. *Developmental Neuropsychology*, *34*(1), 66–84. https://doi.org/10.1080/87565640802564648

Raven, J. (1998). *Raven's progressive matrices*. Oxford: OUP.

Raven, J. (2000). The Raven's Progressive Matrices: Change and Stability over Culture and Time. *Cognitive Psychology*, *41*(1), 1–48. https://doi.org/10.1006/cogp.1999.0735

Redmond, S. M., Thompson, H. L., & Goldstein, S. (2011). Psycholinguistic profiling differentiates specific language impairment from typical development and from attention-deficit/hyperactivity disorder. *Journal of Speech, Language, and Hearing Research*, *54*(1), 99–117.

Reilly, S., Tomblin, B., Law, J., McKean, C., Mensah, F. K., Morgan, A., Wake, M. (2014). Specific language impairment: a convenient label for whom?: SLI: a convenient label for whom? *International Journal of Language & Communication Disorders*, *49*(4), 416–451. https://doi.org/10.1111/1460-6984.12102

Riches, N.G., Loucas, T., Baird, G., Charman, T., & Simonoff, E. (2011). Non-word repetition in adolescents with Specific Language Impairment and Autism plus Language Impairments: A qualitative analysis. *Journal of Communication Disorders*, *44*(1), 23–36. https://doi.org/10.1016/j.jcomdis.2010.06.003

Riches, Nick 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.

Risi, S., Lord, C., Gotham, K., Corsello, C., Chrysler, C., Szatmari, P., Pickles, A. (2006). Combining Information From Multiple Sources in the Diagnosis of Autism Spectrum Disorders. *Journal of the American Academy of Child & Adolescent Psychiatry*, *45*(9), 1094–1103. https://doi.org/10.1097/01.chi.0000227880.42780.0e

Rispens, J., & Parigger, E. (2010). Non-word repetition in Dutch-speaking children with specific language impairment with and without reading problems. *British Journal of Developmental Psychology*, *28*(1), 177–188. https://doi.org/10.1348/026151009X482633

Rizzi, L. (1990). *Relativized minimality*. The MIT Press.

Roberts, J. A., Rice, M. L., & Tager-Flusberg, H. (2004). Tense marking in children with autism. *Applied Psycholinguistics*, *25*(3), 429–448.

Rogers, S. J. (1996). Brief report: Early intervention in autism. *Journal of Autism and Developmental Disorders*, *26*(2), 243–246. https://doi.org/10.1007/BF02172020

Roid, G. H., & Sampers, J. L. (2004). *Merrill-Palmer-revised: scales of development.* IL: Stoelting.

Ronald, A., & Hoekstra, R. A. (2011). Autism Spectrum Disorders and Autistic Traits: A Decade of New Twins Studies. *American Journal of Medical Genetics Part B*, (156), 255–274.

Rossetti, L. M. (2006). *The Rossetti infant-toddler language scale*. East Moline, IL: LinguiSystems.

Samson, F., Mottron, L., Soulières, I., & Zeffiro, T. A. (2012). Enhanced visual functioning in autism: An ALE meta-analysis. *Human Brain Mapping*, *33*(7), 1553–1581. https://doi.org/10.1002/hbm.21307

Schaeffer, J. (2016). Linguistic and Cognitive Abilities in Children with Specific Language Impairment as Compared to Children with High-Functioning Autism. *Language Acquisition*, 1–19. https://doi.org/10.1080/10489223.2016.1188928

Schaeffer, J. (2017). Are Children with High-Functioning Autism Better at Syntax than Typically Developing Children? The Case of Dutch Object Relative Clauses. In M. LaMendola and J. Scott, (Eds.) *Proceedings of the 41th Annual Boston University Conference on Language Development. Volume 2: [...]*. Somerville, MA: Cascadilla Press, pp. 576-587.

Scheidnes, M., & Tuller, L. (2013). L2 Children Embed Normally, but Children with SLI Do Not. In M. Lobo & A. L. Santos (Eds.) *Selected Proceedings of the Romance Turn V*. Newcastle upon Tyne, UK: Cambridge Scholars Publishing, pp 261-284.

Schneider, P., Dubé, R. V., & Hayward, D. (2005). The Edmonton Narrative Norms Instrument. University of Alberta Faculty of Rehabilitation Medicine website.

Schoen, E., Paul, R., & Chawarska, K. (2011). Phonology and vocal behavior in toddlers with autism spectrum disorders. *Autism Research*, *4*(3), 177–188. https://doi.org/10.1002/aur.183

Schopler, E., Reichler, R. J., & Renner, B. R. (2002). *The Childhood Autism Rating Scale (CARS).* Los Angeles, CA: Western Psychological Services.

Scrucca, L., Fop, M., Murphy, T. B., & Raftery, A. E. (2016). mclust 5: Clustering, Classification and Density Estimation Using Gaussian Finite Mixture Models. *The R Journal* (*8)1*, 289-317.

Seeff-Gabriel, B., Chiat, S., & Dodd, B. (2010). Sentence imitation as a tool in identifying expressive morphosyntactic difficulties in children with severe speech difficulties. *International Journal of Language & Communication Disorders*, *45*(6), 691–702. https://doi.org/10.3109/13682820903509432

Semel, E. ., Wiig, E. H., & Secord, W. (1995). *CELF-III: Clinical Evaluation of Language Functions: Diagnostic Battery*. The Psychological Corporation, CA: Harcourt Brace & Co.

Shriberg, L. D., & Lohmeier, H. L. (2008). *The Syllable Repetition Task (SRT)*. Phonology Project, Waisman Center, University of Wisconsin-Madison, WI, USA.

Shriberg, L. D., Paul, R., McSweeny, J. L., Klin, A., Cohen, D. J., & Volkmar, F. R. (2001). Speech and prosody characteristics of adolescents and adults with high-functioning autism and Asperger syndrome. *Journal of Speech, Language, and Hearing Research*, *44*(5), 1097–1115.

Siegel, D. J., Minshew, N. J., & Goldstein, G. (1996). Wechsler IQ profiles in diagnosis of high-functioning autism. *Journal of Autism and Developmental Disorders*, *26*(4), 389–406.

Siller, M., & Sigman, M. (2008). Modeling longitudinal change in the language abilities of children with autism: Parent behaviors and child characteristics as predictors of change. *Developmental Psychology*, *44*(6), 1691–1704. https://doi.org/10.1037/a0013771

Silleresi, S., Tuller, L., Delage, H., Durrleman, S., Bonnet-Brilhault, F., Malvy, J., & Prevost, P. (in press). Sentence repetition and language impairment in French-speaking children with ASD. In A. Gavarro (Ed.) *Language Acquisition and Language Disorders*, Amsterdam: John Benjamins, pp. 235–258.

Smith, N., & Tsimpli, I. M. (1995). *The Mind of a Savant: Language, Learning and Modularity*. Oxford, UK and Cambridge, MA: Basil Blackwell.

Smolík, F., & Vávru, P. (2014). Sentence Imitation as a Marker of SLI in Czech: Disproportionate Impairment of Verbs and Clitics. *Journal of Speech Language and Hearing Research*, 837. https://doi.org/10.1044/2014_JSLHR-L-12-0384

Snow, R. E., Kyllonen, P. C., & Marshalek, B. (1984). The topography of ability and learning correlations. In R. J. Sternberg (Ed.) *Advances in the Psychology of Human Intelligence*, Vols. 1–2. Hillsdale, NJ: Erlbaum, pp. 47–103.

Soni, K. G., & Patel, A. (2017). Comparative Analysis of K-means and K-medoids Algorithm on IRIS Data. *International Journal of Computational Intelligence Research*, *13*(5), 899–906.

Soulières, I., Dawson, M., Samson, F., Barbeau, E. B., Sahyoun, C. P., Strangman, G. E., Mottron, L. (2009). Enhanced visual processing contributes to matrix reasoning in autism. *Human Brain Mapping*, *30*(12), 4082–4107. https://doi.org/10.1002/hbm.20831

Stevenson, J. L., & Gernsbacher, M. A. (2013). Abstract Spatial Reasoning as an Autistic Strength. *PLoS ONE*, *8*(3), e59329. https://doi.org/10.1371/journal.pone.0059329

Stokes, S. F., Wong, A. M., Fletcher, P., & Leonard, L. B. (2006). Nonword repetition and sentence repetition as clinical markers of specific language impairment: The case of Cantonese. *Journal of Speech, Language, and Hearing Research*, *49*(2), 219–236.

Su, Y., 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. https://doi.org/10.1016/j.rasd.2014.07.008

Sukenik, N. (2017). *Sémantique lexicale et profils langagiers d'enfants avec autisme de langue hébraïque*. Université de Tours.

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*. https://doi.org/10.3389/fpsyg.2018.00279

Tager-Flusberg, H., & Joseph, R. M. (2003). Identifying neurocognitive phenotypes in autism. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *358*(1430), 303–314. https://doi.org/10.1098/rstb.2002.1198

Tager-Flusberg, H. (1981). On the nature of linguistic functioning in early infantile autism. *Journal of Autism and Developmental Disorders*, *11*(1), 45–56.

Tager-Flusberg, H. (2000). The challenge of studying language development in children with autism. *Methods for Studying Language Production*, 313–332.

Tager-Flusberg, H. (2006). Defining language phenotypes in autism. *Clinical Neuroscience Research*, *6*(3–4), 219–224. https://doi.org/10.1016/j.cnr.2006.06.007

Tager-Flusberg, H. (2015). Defining language impairments in a subgroup of children with autism spectrum disorder. *Science China Life Sciences*, *58*(10), 1044–1052. https://doi.org/10.1007/s11427-012-4297-8

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), 1–21.

Tager-Flusberg, H., & Kasari, C. (2013). Minimally Verbal School-Aged Children with Autism Spectrum Disorder: The Neglected End of the Spectrum: Minimally verbal children with ASD. *Autism Research*, *6*(6), 468–478. https://doi.org/10.1002/aur.1329

Tager-Flusberg, H., Paul, R., Lord, C., Volkmar, F., Paul, R., & Klin, A. (2005). Language and communication in autism. *Handbook of Autism and Pervasive Developmental Disorders*, *1*, 335–364.

Talkowski, M. E., Minikel, E. V., & Gusella, J. F. (2014). Autism Spectrum Disorder Genetics: Diverse Genes with Diverse Clinical Outcomes. *Harvard Review of Psychiatry*, 65–75. https://doi.org/10.1097/HRP.0000000000000002

Taylor, L. J., Maybery, M. T., Grayndler, L., & Whitehouse, A. J. O. (2014). Evidence for Distinct Cognitive Profiles in Autism Spectrum Disorders and Specific Language Impairment. *Journal of Autism and Developmental Disorders*, *44*(1), 19–30. https://doi.org/10.1007/s10803-013-1847-2

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. https://doi.org/10.1007/s10803-013-1853-4

Terzi, A., Marinis, T., Kotsopoulou, A., & Francis, K. (2014). Grammatical abilities of Greek-speaking children with autism. *Language Acquisition*, *21*(1), 4–44.

Terzi, A., Zafeiri, A., Marinis, T., & Francis, K. (2017). Object Clitics in the Narratives of High-Functioning Children with Autism. In M. LaMendola & J. Scott (Eds.) *Proceedings of the 41st annual Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press, pp. 637–650.

Theodorou, E., Kambanaros, M., & Grohmann, K. K. (2017). Sentence Repetition as a Tool for Screening Morphosyntactic Abilities of Bilectal Children with SLI. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.02104

Thordardottir, E., Kehayia, E., Mazer, B., Lessard, N., Majnemer, A., Sutton, A., Chilingaryan, G. (2011). Sensitivity and Specificity of French Language and Processing Measures for the Identification of Primary Language Impairment at Age 5. *Journal of Speech Language and Hearing Research*, *54*(2), 580. https://doi.org/10.1044/1092-4388(2010/09-0196)

Thordardottir, E., Rothenberg, A., Rivard, M.-E., & Naves, R. (2006). Bilingual assessment: Can overall proficiency be estimated from separate measurement of two languages? *Journal of Multilingual Communication Disorders*, *4*(1), 1–21. https://doi.org/10.1080/14769670500215647

TNS Opinion & Social at the request of Directorate-General Education and Culture, Directorate-General for Translation and Directorate-General for Interpretation. (2012). *Europeans and their languages*. European Commission.

Tomblin, B. (2011). Co-morbidity of autism and SLI: kinds, kin and complexity: Co-morbidity of autism and SLI: kinds, kin and complexity. *International Journal of Language & Communication Disorders*, *46*(2), 127–137. https://doi.org/10.1111/j.1460-6984.2011.00017.x

TopbaŞ, S., Kaçar-Kütükçü, D., & Kopkalli-Yavuz, H. (2014). Performance of children on the Turkish Nonword Repetition Test: Effect of word similarity, word length, and scoring. *Clinical Linguistics & Phonetics*, *28*(7–8), 602–616. https://doi.org/10.3109/02699206.2014.927003

Tovar, A. T., Fein, D., & Naigles, L. R. (2015). Grammatical Aspect Is a Strength in the Language Comprehension of Young Children With Autism Spectrum Disorder. *Journal of Speech Language and Hearing Research*, *58*(2), 301. https://doi.org/10.1044/2014_JSLHR-L-13-0257

Tsimpli, I. M., Peristeri, E., & Andreou, M. (2016). Narrative production in monolingual and bilingual children with specific language impairment. *Applied Psycholinguistics*, *37*(01), 195–216. https://doi.org/10.1017/S0142716415000478

Tuller, L. (2015). Clinical use of parental questionnaires in multilingual contexts. In S. Armon-Lotem, J. de Jong, N. Meir (Eds.) *Methods for assessing multilingual children: Disentangling bilingualism from Language Impairment*, Vol. Multilingual Matters, Bristol: UK, pp. 301–330.

Tuller, L., Abboud, L., Ferré, S., Fleckstein, A., Prevost, P., dos Santos, C., Zebib, R. (2015). Specific language impairment and bilingualism: assembling the piece. In C. Hamann and E. Ruigendijk (Eds) *Language Acquisition and Development: Proceedings of GALA 2013*. Newcastle: Cambridge Scholars. pp. 533– 567.

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.R. Naigles (Ed.) *Language and the human lifespan series. Innovative* *investigations of language in autism spectrum disorder*. Berlin: de Gruyter, pp. 115–140.

Tuller, L., Hamann, C., Chilla, S., Ferré, S., Morin, E., Prevost, P., Zebib, R. (2018). Identifying language impairment in bilingual children in France and in Germany: Language impairment in bilingual children. *International Journal of Language & Communication Disorders*, *53*(4), 888–904. https://doi.org/10.1111/1460-6984.12397

Turner, L. M., Stone, W. L., Pozdol, S. L., & Coonrod, E. E. (2006). Follow-up of children with autism spectrum disorders from age 2 to age 9. *Autism*, *10*(3), 243–265. https://doi.org/10.1177/1362361306063296

Ullman, M. T. (2004). Contributions of memory circuits to language: the declarative/procedural model. *Cognition*, *92*(1–2), 231–270. https://doi.org/10.1016/j.cognition.2003.10.008

Valicenti-McDermott, M., Tarshis, N., Schouls, M., Galdston, M., Hottinger, K., Seijo, R., Shinnar, S. (2013). Language Differences Between Monolingual English and Bilingual English-Spanish Young Children With Autism Spectrum Disorders. *Journal of Child Neurology*, *28*(7), 945–948. https://doi.org/10.1177/0883073812453204

van Bakel, M. M. E., Delobel-Ayoub, M., Cans, C., Assouline, B., Jouk, P.-S., Raynaud, J.-P., & Arnaud, C. (2015). Low but Increasing Prevalence of Autism Spectrum Disorders in a French Area from Register-Based Data. *Journal of Autism and Developmental Disorders*, *45*(10), 3255–3261. https://doi.org/10.1007/s10803-015-2486-6

Volkmar, F., Chawarska, K., & Klin, A. (2005). Autism in Infancy and Early Childhood. *Annual Review of Psychology*, *56*(1), 315–336. https://doi.org/10.1146/annurev.psych.56.091103.070159

Walenski, M., Mostofsky, S. H., Gidley-Larson, J. C., & Ullman, M. T. (2008). Brief Report: Enhanced Picture Naming in Autism. *Journal of Autism and Developmental Disorders*, *38*(7), 1395–1399. https://doi.org/10.1007/s10803-007-0513-y

Walenski, M., Tager-Flusberg, H., & Ullman, M. (2006). Language in Autism. In S. Moldin & J. Rubenstein (Eds.) *Understanding Autism*. CRC Press, pp. 175–203. https://doi.org/10.1201/9781420004205.ch9

Waterhouse, L., & Fein, D. (1982). Language skills in developmentally disabled children. *Brain and Language*, *15*(2), 307–333.

Weitlauf, A. S., Gotham, K. O., Vehorn, A. C., & Warren, Z. E. (2014). Brief Report: DSM-5 “Levels of Support:” A Comment on Discrepant Conceptualizations of Severity in ASD. *Journal of Autism and Developmental Disorders*, *44*(2), 471–476. https://doi.org/10.1007/s10803-013-1882-z

Whitehouse, A. J. O., Barry, J. G., & Bishop, D. V. M. (2008). Further defining the language impairment of autism: Is there a specific language impairment subtype? *Journal of Communication Disorders*, *41*(4), 319–336. https://doi.org/10.1016/j.jcomdis.2008.01.002

Whitehouse, A. J. O., Maybery, M. T., & Durkin, K. (2007). Evidence against poor semantic encoding in individuals with autism. *Autism*, *11*(3), 241–254. https://doi.org/10.1177/1362361307076860

WHO. (2018). ICD-11. Retrieved from https://icd.who.int

Wiggins, L. D., Robins, D. L., Adamson, L. B., Bakeman, R., & Henrich, C. C. (2012). Support for a Dimensional View of Autism Spectrum Disorders in Toddlers. *Journal of Autism and Developmental Disorders*, *42*(2), 191–200. https://doi.org/10.1007/s10803-011-1230-0

Williams, D., Botting, N., & Boucher, J. (2008). Language in autism and specific language impairment: Where are the links? *Psychological Bulletin*, *134*(6), 944–963. https://doi.org/10.1037/a0013743

Williams, D., Payne, H., & Marshall, C. (2013). Non-word Repetition Impairment in Autism and Specific Language Impairment: Evidence for Distinct Underlying Cognitive Causes. *Journal of Autism and Developmental Disorders*, *43*(2), 404–417. https://doi.org/10.1007/s10803-012-1579-8

Williams, P. E., Weiss, L. G., & Rolfhus, E. (2003). *WISC-IV Technical Report. Theoretical Model and Test Blueprint* (No. 1). The Psychological Corporation; London: Pearson.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraircing function of wrong beliefs in young children's understanding of deception. *Cognition*, pp. 103–128.

Wittke, K., Mastergeorge, A. M., Ozonoff, S., Rogers, S. J., & Naigles, L. R. (2017). Grammatical Language Impairment in Autism Spectrum Disorder: Exploring Language Phenotypes Beyond Standardized Testing. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.00532

Wodka, E. L., Mathy, P., & Kalb, L. (2013). Predictors of Phrase and Fluent Speech in Children With Autism and Severe Language Delay. *Pediatrics*, *131*(4), 1128–1134. https://doi.org/10.1542/peds.2012-2221

Wolk, L., & Edwards, M. L. (1993). The emerging phonological system of an autistic child. *Journal of Communication Disorders*, *26*(3), 161–177.

Young, E. C., Diehl, J. J., Morris, D., Hyman, S. L., & Bennetto, L. (2005). The use of two language tests to identify pragmatic language problems in children with autism spectrum disorders. *Language, Speech, and Hearing Services in Schools*, *36*(1), 62–72.

Zebib, R., Henry, G., Khomsi, A., Messarra, C., & Hreich, E. (2017). *Evaluation du Langage Oral chez l'enfant Libanais: ELO-L*. Antelias: Liban Tests Editions.

Zebib, R., Tuller, L., Prévost, P., & Morin, E. (2013). Formal Language Impairment in French-sèpeaking children with ASD: a comparative ASD/SLI study. In *Advances in Language Acquisition*. Newcastle: Cambridge Scholars Publisher.

Zhou, P., Crain, S., Gao, L., Tang, Y., & Jia, M. (2015). The Use of Grammatical Morphemes by Mandarin-Speaking Children with High Functioning Autism. *Journal of Autism* *and Developmental Disorders*, *45*(5), 1428–1436. https://doi.org/10.1007/s10803-014-2304-6

Zimmerman, I., Steiner, V., & Pond, R. (2002). *Preschool Language Scale, Fourth Edition*. San Antonio, TX: The Psychological Corporation.

Appendix

APPENDIX 1. THE LITMUS TASKS

LITMUS-NWR task

Control items			
CCV	[plu]	CVC	[kip], [fuk], [kis],
sCV	[spu]		
Low complexity items			
CV.CCV	[paklu], [fupli]	CCV.CV	[plifu]
CV.CVC	[pukif], [kafip]	CV.CVs	[fapus]
CV.CV.CV	[kifapu], [pufaki]		
Average complexity items			
CCV.CVC	[flukif], [klifak]	CCV.CCV	[flaplu], [plaklu]
CCV.CV.CV	[flipuka], [klipafu]	CV.CV.CCV	[kupifla]
CV.CCV.CV	[kuflapi], [piklafu]	CV.CV.CVC	[kapufik], [pifakup]
CV.CV.CVs	[pifukas]		
CCVL	[plal], [klil]		
CVCs	[fips], [piks]	CVsC	[pusk], [kusp]
CCVs	[flis], [klis]	sCV.CV	[skafu]
sCCV	[skla], [spli]	sCVC	[skap], [spaf]
CVL.CV	[pilfu],[filpa]	CVs.CV	[kusp]
High complexity items			
CCVCs	[pliks], [klups]	CCVsC	[plusk]
CV.CVL.CV	[kufalpi], [kupalfi]	CV.CVs.CV	[pafuski], [fikuspa]
sCV.CV.CV	[skapufi], [spakifu]		

LITMUS-SR task

	Sentences	English glosses	English translation
	SVO Simple Present		
3S	La maîtresse punit les enfants.	the-fem-sg teacher punishes the-pl children	The teacher punishes the children
3S	Le garçonprend un bain.	the-masc-sg boy takes a-masc bath	The boy takes a bath
3S	La maman lit une histoire.	the-fem-sg mother reads a-fem story	The mother reads a story
3P	Les chats boivent du lait	the-pl cats drink-3pl some-masc milk	The cats drink some milk
3P	Les parents punissent les enfants.	the-pl parents punish-3pl the-pl children	The parents punish the children
3P	Les enfants prennent un bain.	the-pl children take-3pl a-masc bath	The children take a bath
	SVO Past tense		
3S	La maman a fermé la fenêtre.	the-fem mother has opened the window	The mother opened the window
3S	Le lapin a mangé la carotte.	the-masc-sg rabbit has eaten the-fem-sg carrot	The rabbit ate the carrot
3S	Le singe a pris la banane.	the-masc-sg monkey has taken the-fem banana	The monkey took the banana
3P	Les enfants ont fermé la porte.	the-pl children have-3pl opened the door	The children opened the door
3P	Les parents ont rangé les jouets.	the-pl parents have-3pl put+away the-pl toys	The parents put away the toys
3P	Les tortues ont mangé la salade.	the-pl turtles have-3pl eaten the-fem lettuce	The turtles ate the lettuce
	Wh-Questions		
Qui wh-object	Qui le monsieur regarde ?	who the-masc-sg man looks+at	Who is the man looking at?
Qui wh-object	Qui la maîtresse punit ?	who the-fem teacher punishes	Who is the teacher punishing?
Qui wh-object	Qui la mamie connait ?	who the-fem grandma knows	Who does the grandma know?
Quel wh-object	Quel garçon le monsieur dessine ?	which boy the-masc-sg man draws	Which boy is the man drawing?
Quel wh-object	Quel enfant la maîtresse punit?	which child the-fem teacher punishes	Which child is the teacher punishing?
Quel wh-object	Quel garçon le papy connaît ?	which boy the-masc-sg grandpa knows	Which boy does the grandpa know?

	Argument Clauses		
Finite	Le garçon dit que la maman a lu un livre.	the-masc-sg boy says that the-fem mommy has read a-masc book	The boy says that the mommy read a book
Finite	La dame dit que le garçon a pris le ballon.	the-fem-sg mommy says that the-masc-sg boy has taken the-sg-masc ball	The mommy says that the boys took the ball
Finite	La fille croit que le papi a cassé un verre.	the-fem-sg girl believes that the-masc-sg grandpa has broken a-masc glass	The girl believes that the grandpa broke a glass
Nonfinite	Le lapin veut manger la salade maintenant.	the-masc-sg rabbit like to+drive the-fem car now	The rabbit now wants to eat the lettuce
Nonfinite	Le papa sait très bien conduire la voiture	the-masc-sg daddy knows very well to+drive the-fem car	The daddy knows how to drive the car very well
Nonfinite	La maman sait très bien dessiner des lapins	the-fem-sg mommy knows very well to+draw some-pl rabbits	The mommy knows how to draw rabbits very well
	Relative Clauses		
Subject relative	Tu vois le garçon qui a dessiné la mamie.	You see the boy-masc-sg that-subject has drawn the-fem grandma	You see the boys who drew the mommy
Subject relative	J'ai vu le chat qui a griffé la vache.	I have seen the-masc-sg cat that-subject has scratched the-fem cow	I saw the cat that scratched the cow
Subject relative	J'ai vu le chien qui a mordu le cheval.	I have seen the-masc dog that-subject has bitten the-masc horse	I saw the dog that bit the horse
Object relative	Je vois le garçon que la fille a poussé.	I see the-masc-sg boy that-object the-fem girl has pushed	I see the boy that the girl pushed
Object relative	Tu as vu la vache que le chat a griffée.	You have seen the-fem cow that-object the-masc-sg cat has scratched	Yousaw the cow that the cat scratched
Object relative	Tu as vu le cheval que le chien a mordu.	You have seen the-masc-sg horse that-object the-masc dog has bitten	You saw the horse that the dog bit

APPENDIX 2: PABIQ QUESTIONNAIRE

Nom de l'enfant : …………………..   Date de l'entretien ……………………….

Nom de l'orthophoniste ayant conduit l'entretien : ……………………….   Entretien avec mère ▫ père ▫

Tierce personne ayant servi d'intermédiaire avec parent OUI ou NON. Lien avec la famille ………………

****

PaBiQ (Questionnaire pour parents d'enfants bilingues)[1]

1. Informations générales sur l'enfant

1.1 Date de naissance : ____________ Age actuel : ________ Pays de naissance : _________________________
années ; mois

1.2 Si le lieu de naissance est différent du pays de résidence actuelle, date d'arrivée en France : ________________
(mois/année)

1.3 Quelles langues est-ce que votre enfant parle actuellement (plus que des mots isolés) ?

langue (à preciser)………………….	français	autre ……………

1.4 Selon vous, dans quelle langue se sent-il le plus à l'aise ? ______________________________________

2. Histoire précoce de l'enfant : langage et autre

2.1 A quel âge votre enfant a produit son premier mot (Ex: Papa, lolo, encore, pain) ?
15 mois ou plus jeune ▫ entre 16 et 24 mois ▫ 25 mois ou plus vieux▫ ou _____________ âge

2.2 Ces premiers mots étaient en quelle langue ? ______________________

2.2 Vers quel âge votre enfant a mis ensemble des mots pour faire de petites phrases (Ex: encore pain, a plus gâteau)?

24 mois ou plus jeune ▫ entre 25 et 30 mois ▫ 31 mois ou plus vieux▫ ou _____________ âge

2.3 Avant l'âge de 3-4 ans de votre enfant, vous êtes-vous inquiété du langage de votre enfant ? OUI ou NON

2.4 Votre enfant a-t-il eu des problèmes d'audition ou des otites fréquents ? OUI ou NON

2.5 De façon générale, avant l'âge de 4 ans, votre enfants a-t-il été en contact avec **:**

	0 Jamais	**1** Rarement/de temps en temps	**2** La moitié du temps	**3** Souvent	**4** Très Souvent/ Toujours
langue (à preciser) ……………………					
français					
autre					

[1]Ce questionnaire est une traduction/adaptation du Questionnaire for Parents of Bilingual Children (COST Action IS0804, 2011). Il s'agit de la version courte d'un questionnaire plus long piloté par des groupes de recherche dans plusieurs pays de l'Action COST IS0804, qui a été en partie basé sur l'ALEQ (Paradis, 2011) et sur l'ALDeQ (Paradis et al., 2010). Référence complète : COST Action IS0804 (2011). Questionnaire for Parents of Bilingual Children (PABIQ). http://www.bi-sli.org.

1

2.6 Dans quels contextes votre enfant a-t-il eu des contacts ces langues, avant l'âge de 4 ans ? Je vais vous donner plusieurs exemples : à chaque fois, vous devrez me dire si oui ou non votre enfant a eu des échanges avec ces personnes ou dans ces situations avant 4 ans. Si vous me répondez oui, je vais vous demander de préciser dans quelle langue, et à partir de quel âge exactement.

Avant l'âge de 4 ans, votre enfant a-t-il été en contact avec…		**Si OUI, dans quelle langue ? A partir de quel âge ?**					
		Langue ……………		Français		autre	
Oui/non		Oui/non	Age	Oui/non	Age	Oui/non	Age
Sa mère							
Son père							
Ses grands-parents							
La nounou							
d'autres adultes ……………							
Ses frères et sœurs							
Garderie/crèche							
Ecole maternelle							
Age du 1^er contact							
Nombre total de contextes de contact		T			A	B	C
	Taux d'exposition à chaque langue				A……… / T……..	B…… / T……	C…… / T…..

3. Habiletés actuelles

	langue ……………	français	autre
3.1 Comparé à d'autres enfants du même âge, comment pensez-vous que votre enfant s'exprime en … ? *0 = pas très bien/ pas aussi bien qu'eux ; 1 = un peu moins bien / à peu près comme eux avec quelques différences ; 2 = pareil/ en général oui ; 3 = très bien, mieux, beaucoup mieux*	0 1 2 3	0 1 2 3	0 1 2 3
3.2 Comparé à d'autres enfants du même âge, est-ce que votre enfant a des difficultés pour produire des phrases correctes ? *0 = beaucoup de difficultés ; 1 = quelques difficultés ; 2 = pareil/ en général ; 3 = pas de difficultés, mieux que d'autres enfants*	0 1 2 3	0 1 2 3	0 1 2 3

3.3 Voyez-vous une différence entre le niveau de votre enfant dans les deux langues ? *0 = beaucoup de différence ; 1 = quelques différences ; 2 = pareil/ en général ; 3 = pas de différence*	0 1 2 3	0 1 2 3	0 1 2 3
3.4 Est-ce que votre enfant se sent frustré (ou mécontent, agacé) quand il ne peut pas communiquer en … ? *0 = très frustré/presque toujours frustré/très souvent frustré; 1 = souvent frustré/oui ; 2 = quelquefois frustré mais pas souvent ; 3 = jamais/presque jamais frustré/non*	0 1 2 3	0 1 2 3	0 1 2 3
3.5 Est-ce que vous pensez que votre enfant parle … comme un enfant …. du même âge qui ne parle que le ….. ? *0 = pas très bien/ pas aussi bien qu'eux ; 1 = un peu moins bien / à peu près comme eux avec quelques différences ; 2 = pareil/ en général oui ; 3 = très bien, mieux, beaucoup mieux*	0 1 2 3	0 1 2 3	0 1 2 3
TOTAL Habiletés actuelles par langue	***/15***	***/15***	***/15***

4. Quelles langues sont utilisées au sein de la famille ?

4.1 Quelle langue est utilisée entre vous (la mère/ le père/ autre adulte/ frères et sœurs) et votre enfant ?
0 = jamais ; 2 = rarement/de temps en temps ; 2 = la moitié du temps; 3 = souvent ; 4= très souvent/toujours

	Langue ………	**Français**	**Autre**
Mère ↔ enfant	0 1 2 3 4	0 1 2 3 4	0 1 2 3 4
Père ↔ enfant	0 1 2 3 4	0 1 2 3 4	0 1 2 3 4
Autre adulte ↔ enfant	0 1 2 3 4	0 1 2 3 4	0 1 2 3 4
frères et sœurs ↔ enfant	0 1 2 3 4	0 1 2 3 4	0 1 2 3 4
Total utilisation de chaque langue	/16	/16	/16

5. Langues parlées dans d'autres contextes

5.1 Quelle(s) langue(s) est utilisée dans les situations suivantes ?
0 = jamais ; 2 = rarement/de temps en temps ; 2 = la moitié du temps; 3 = souvent ; 4= très souvent/toujours

	Langue ………	**Français**	**Autre**
Echanges			
enfant ↔ camarades	0 1 2 3 4	0 1 2 3 4	0 1 2 3 4
enfant ↔ amis de famille venant régulièrement	0 1 2 3 4	0 1 2 3 4	0 1 2 3 4
Activités			
Lire (livres, revus, BD, journal)	0 1 2 3 4	0 1 2 3 4	0 1 2 3 4
Regarder la télévision/ film ; utiliser l'ordinateur	0 1 2 3 4	0 1 2 3 4	0 1 2 3 4
Raconter des événements ou des histoires	0 1 2 3 4	0 1 2 3 4	0 1 2 3 4

	langue …………	français	autre
Total utilisation de chaque langue dans d'autres contextes (Richesse linguistique) :	**/14**	**/14**	**/14**

5.2 Si votre enfant parle tout seule quand il joue, c'est dans quelle langue ? ____________________________

6. Informations sur la mère et le père

6.1 Informations sur la mère

6.1.1 Dans quel pays êtes-vous née ?____

6.1.2 Si vous exercez une profession actuellement, quelle est la langue que vous utilisez dans votre travail ? ________

6.1.3 Quelles sont les langues que vous maîtrisez et à quel niveau ?

	0 Inexistant (maximum quelques mots)	1 Faible (se débrouille difficilement)	2 Moyen (se débrouille)	3 Bon (est à l'aise)	4 Excellent (parle couramment)
langue ……………					
français					
Autre langue ………….					

6.2 Informations sur le père

6.2.1 Dans quel pays êtes-vous née ?____

6.2.2 Si vous exercez une profession actuellement, quelle est la langue que vous utilisez dans votre travail ? ________

6.2. 3 Quelles sont les langues que vous maîtrisez et à quel niveau ?

	0 Inexistant (maximum quelques mots)	1 Faible (se débrouille difficilement)	2 Moyen (se débrouille)	3 Bon (est à l'aise)	4 Excellent (parle couramment)
langue ……………					
français					
Autre langue ………….					

Appendix 3. Comparisons between MO-ASD-LN and BI-ASD-LN groups on SR and NWR tasks

Results on SR

Groups		Total score	Present tense 3P	Present tense 3S	Past Tense 3P	Past Tense 3S	Wh-object quel	Wh-object qui	Relative clauses OR	Relative clauses SR	Argument clauses finite	Argument clauses nonfinite
		Identical repetition										
MO-ASD-LN with high NVIQ (n = 8) / BI-ASD-LN with high NVIQ (n =3)	U	10	3	4	3	2.5	4	3	3	3.5	2	2
	p	.656	.593	1.000	.593	.487	1.000	.593	.596	.724	.386	.386
MO-ASD-LN with low NVIQ (n = 4) / BI-ASD-LN with low NVIQ (n = 1)	U	.000	2	.000	2	.000	1.5	1.5	1.5	1.000	.000	.500
	p	.147	1.000	.400	1.000	.136	.617	.617	.709	.414	.114	.264
MO-ASD-LN with average NVIQ (n = 11) / BI-ASD-LN (n = 1)	U	4.5	4.5	5.5	2	2	5.5	4.5	4.5	4.5	5.5	.000
	p	.771	.771	1.000	.237	.237	1.000	.771	.771	.771	1.000	1.000
		Target Structure										
MO-ASD-LN with high NVIQ (n = 8) / BI-ASD-LN with high NVIQ (n =3)	U	11.5	3.5	4	3.5	3.54	4	3.5	3	4	4	3.5
	p	.906	.593	1.000	.724	.724	1.000	.724	.593	1.000	1.000	.724
MO-ASD-LN with low NVIQ (n = 4) / BI-ASD-LN with low NVIQ (n = 1)	U	2	2	2	2	2	1.5	1.5	1.5	1.5	2	2
	p	1.000	1.000	1.000	1.000	1.000	.617	.617	.709	.617	1.000	1.000
MO-ASD-LN with average NVIQ (n = 11) / BI-ASD-LN (n = 1)	U	4.5	5.5	5.5	5.5	5.5	5.5	2	.000	2	2	.000
	p	.771	1.000	1.000	1.000	1.000	1.000	.237	1.000	.237	.237	1.000

Results on NWR

		Total score	1-syll	2-syll	3-syll	0-clus	1-clus	2-clus
MO-ASD-LN with high NVIQ (n = 8) / BI-ASD-LN with high NVIQ (n =3)	*U*	8.5	8	8	8	8	8	8
	p	.469	.401	.408	.401	.408	.408	.408
MO-ASD-LN with low NVIQ (n = 5) / BI-ASD-LN with low NVIQ (n = 1)	*U*	.000	.000	.000	.000	.000	.000	.000
	p	.132	.136	.147	.114	.132	.117	.114
MO-ASD-LN with average NVIQ (n = 12) / BI-ASD-LN (n = 3)	*U*	16.5	12	12	12	12	12	12
	p	.826	.382	.378	.382	.382	.378	.382