Same, but different?

Early language characteristics in children with autism spectrum disorders, compared to children with language impairment and children with typical development

Stian Barbo Valand

Master Thesis in Special Needs Education
Department of Special Needs Education
Faculty of Educational Sciences

UNIVERSITY OF OSLO

Spring 2013

Early language characteristics in children with autism spectrum disorders, compared to children with language impairment and children with typical development

2013

Same, but different? Early language characteristics in children with autism spectrum disorders, compared to children with language impairment and children with typical development.

Stian Barbo Valand

http://www.duo.uio.no/

Printed: Reprosentralen, Universitetet i Oslo

Abstract

Aim: The current study investigated early language characteristics in pre-school children (age 3.0 – 4.11) with autism spectrum disorders (ASD) compared to children with language impairment (LI), and typically developing children, using data from the cross-sectional Autism Birth Cohort Study, at the Norwegian Institute of Public Health. **Background:** There seems to be limited information in the literature regarding language characteristics in such a young sample of children, especially research looking at similarities and differences between children with ASD, LI and typically developing. Particularly looking at young Norwegian children with these difficulties is non-existent. **Method:** The current study is quantitative in methodology. It uses questionnaire and interview data from parental reports of language functioning, and direct testing of cognitive development to compare the three groups. For investigation of group differences analysis of variance (ANOVA) and analysis of covariance (ANCOVA) were used, with the addition of more in-depth investigation of selected items from the main variables. **Results:** The current study shows there were both differences and similarities between the ASD and LI groups on language variables, compared to the TD group which scored significantly better on all variables. However both the ASD and LI group were found to be more similar than expected. **Conclusion:** The potential overlap between ASD and LI on language characteristics might be seen as an indication of overlapping difficulties. However, perhaps the most important potential implication from the current study is better understanding of early language characteristics which may lead to better interventions for both groups.

Acknowledgment

First, I would like to thank my main UiO-supervisor **Ninna Garm** for believing in the project, providing feedback and sharing her insight in the field of special needs education.

I also wish to thank my ABC-supervisor **Synnve Schjølberg** for her dedication, and her interest and knowledge of the field of autism spectrum disorders and language disorders. Not to mention her patience with my SPSS endeavors, including but not limited to navigating the ABC-data file and SPSS syntax editor. It would not have been possible to complete this undertaking without your dedication.

I would also like to thank the **Autism Birth Cohort Study** for providing me with the dataset, and hence making this thesis a possibility. This includes both the ABC **Steering Committee**, with Ezra Susser being especially encouraging, the ABC **Study Group** and **the clinical assessment team** who were my co-workers for many years. Unfortunately you are too many to name.

Additionally I would like to thank **Anne-Siri Øyen** for reading my thesis and providing both valuable comments and correcting my English blunders.

A special thank you goes to **all the children and their families** who dedicated their valuable time to the study. They made this study possible, and hopefully results from the ABC-study will bring greater knowledge about developmental disorders to clinicians and educators, with better and more effective support for those who need extra help.

Stian Barbo Valand

May 2013

Table of contents

List of tables

Main abbreviations

ABC–study: Autism Birth Cohort Study

ADI-R: Autism Diagnostic Interview – Revised

ASD: Autism Spectrum Disorders

CDI: Child Development Inventory

DQ: Developmental Quotient

NIPH: Norwegian Institute of Public Health

FSIQ: Full-scale Intelligence Quotient

LF-ASD: Low-functioning ASD

LI: Language Impairment

Mullen: Mullen Scales of Early Learning

MoBa: The Norwegian Mother and Child Cohort Study

NVIQ: Nonverbal Intelligence Quotient

PPBS: Preschool Play Behavior Scale

SB5: Stanford-Binet Intelligence Scale, 5th edition

TD: Typically Developing

VABS: Vineland Adaptive Behavior Scales

1 Introduction

The following will include a brief introduction to the purpose- and problem statement for the current study, as well as a brief review of the structure of current thesis.

1.1 Purpose and aims of the current study

The purpose of the current study is to investigate expressive language characteristics of three and four year old children with a diagnosis of ASD, compared to children with a primary diagnosis of language impairment (LI) and typically developing children (TD), all drawn from a population based sample. This will be done by comparing expressive language abilities measured across the three groups. Furthermore the study will also explore if there are gender differences for any of the variables under study. Three main research questions were selected for the current study:

1. Are there similarities and/or differences in early language characteristics for the ASD and LI group compared to TD?

2. Are there similarities in social difficulties between the ASD and LI group, compared to the TD group?

3. Are there overlaps in autism symptomatology for the ASD and LI groups, compared to the TD group?

The current study uses data from the Autism Birth Cohort Study (ABC-study). The ABC-study includes a rich dataset meant to investigate many aspects related to ASD and other developmental disorders. Using such a rich dataset provides many unique possibilities. The primary strengths in the current study are the number of participants, and the use of a population based sample. An additional strength is that all the children have been assessed and diagnosed by specialists in clinical psychology and/or child psychiatry. However, there are also limitations in using preexisting data. The main limitation being that all measures in the current study have been pre-selected independent of the aims in this thesis. In this case this means that a set of psychometric and standardized tests, primarily from the field of psychology, have been used. There were no standardized language test used in the ABC-study, but normed questionnaires and interviews of language functions were utilized. An

additional limitation from a speech pathologist perspective is the sole use of the current classification of language disorders from the diagnostic manuals of International Classification of Diseases-10 (ICD-10; World Health Organization, 1992), and Diagnostic and Statistical Manual of Mental Disorders - 4th Edition - Text Revision (DSM-IV-TR; American Psychiatric Association, 2000).

The current study covers the field of medicine, psychology and to lesser extent the science of education. The assessments of language functions provide some details of the children's language abilities, although more specific testing would have been required for a comprehensive assessment of the whole language area. Depending on the readers' theoretical points-of-view, the issues highlighted here may be considered limitations for in- depth knowledge, while for others they may be considered strengths due to the challenge all studies have collecting valid assessment data for such a high number of participants.

1.2 Structure of the thesis

The current chapter includes a brief introduction to the purpose of the current study, as well as the problem statement. Chapter 2 includes the theoretical background for the problem statement. This includes an introduction to typical language development, autism spectrum disorders, language impairment and language characteristics in these two groups separately, as well as an introduction to the potential overlap between these two diagnoses. Chapter 3 includes a description of the method for the current study. This includes description of the participants, instruments, procedures (including how various measures were constructed, and statistical procedures), validity and reliability and lastly ethical considerations. Chapter 4 includes the results presented thematically. Chapter 5 includes a brief recap of the main findings, discussion of these in relation to the problem statement, and also a discussion of additional findings of interest.

2 Theoretical background

The following section will include an introduction to theories of language development, as well as characteristics of early language development. It will also include an overview of autism spectrum disorders, language impairment and social difficulties associated with these.

2.1 Language development

Acquisition of language is an important part of human development. It is a fundamental factor for communication between humans, and starts early in human development. Most infants develop language seemingly effortlessly from babbling at six months, to full sentences by the age of three years (Kuhl, 2004).

According to Owens (2012) language can be understood as a complex and dynamic system of conventional symbols that are used in various modes for thought and communication. This definition emphasizes that language is part of the larger process of communication. Communication can be perceived as the process were participants exchange information, ideas, needs and desires (Owens, 2012). Communication can be accomplished through extra-linguistic (e.g. contextual information) and linguistic codes (e.g. language). This covers aspects of communication stretching from the change of pitch and pauses that alters the meaning of the utterance (supra-segmental devices), to nonverbal use of language in the form of gestures and different linguistic modes (speaking and listening, writing and reading, and signing) (Owens, 2012).

Bloom and Lahey (1978) referenced in Owens (2012) argue that language is made up of different components. They describe three different aspects of language that overlap: content, form and use. The content of language refers to words, and meaning of word combinations (semantics). The form of the language refers to language sounds (phonology), how words are constructed (morphology) and how words are put together to make sentences (syntax). The use of language refers to how language is used and interpreted in social situations (pragmatics). Another similar approach to the one mentioned, is uunderstanding pragmatics as an overall aspect of language, were semantics, morphology, phonology and syntax are "secondary". The context determines the use of communicative form. This is often called the functionalist model (Owens, 2012).

2.1.1 Theories of language development

Many theories from different fields of science try to explain the development of language, from linguistic theory about the construction of language to cognitive sciences about how the brain processes language. No single theory seems to cover all relevant aspects of language development.

According to Bohannon and Bonvillian (2013) there are several theoretical approaches to language acquisition. They describe the classical approaches of behavioral and linguistic theories as outdated by today's understanding, but they survive in revised versions and are foundations for more modern theories. A more up-to-date understanding of language acquisition is proposed through the interactionist position. This position describe four subdivisions, with cognitive-, information processing-, social- and usage/gestural-approaches.

The cognitive approach states that language is just one of many complex cognitive skills, and the process of learning language can be accounted for by many of these abilities. The information processing approach emphasizes that language is learned through decoding stimuli from the environment, which is interpreted and stored in memory for later retrieval. The social approach highlights the social context of language acquisition, and the importance of social interaction. The last of the interactionist positions is the usage/gestural approach, it highlights that language development begins with the use of gestures and develops into the use of vocal signs. This approach also states that human evolution of vocal language can be seen in relation to gestural use in non-human primates. According to this view human evolution started with gestural use, but evolved into vocal language due to needs for warning others in the herd. Bohannon and Bonvillian (2013) describe the various interactionist approaches as eclectic in nature, and the most feasible way to explain the complex nature of language development. It therefore seems correct to understand language as a multimodal discipline, which contains many theories from different fields of science, especially psychology and linguistics.

In summary, it seems to be commonly understood that language development relies on an interaction between human biology and environment. It is also agreed that the child plays an active part in learning language through engaging with the environment. Especially in early development, when the child interacts with its mother and other caregivers using joint attention. In addition, some experiments have shown that acquisition of language happens in

social context only. Children (and even birds) need language models present in order to learn language (or bird song). This means that children do not seem to learn language purely by being exposed to language sounds (Kuhl, 2004).

2.1.2 Characteristics of early typical language development

The following will include a short introduction of typical language development up to pre-school age. This introduction will primarily focus on the developmental milestones in language development, it will to a lesser extent focus on cognitive explanations of these milestones. This is not to undermine the important role of cognitive theories in explaining language development, but it is beyond the scope of this thesis to describe them further. This section is mostly based on Owens (2012) book on language development. Most of the current knowledge is based on acquisition of English language, and it should be noted that some language functions might develop differently in Norwegian. Research on early language development in samples of Norwegian children is few. However, Kristoffersen, Simonsen, Eiesland, and Henriksen (2012) did a study with a sample of 6 500 typically developing, monolingual, Norwegian children from 8 to 36 months. The study collected data using an online version of the MacArthur-Bates Communicative Development Inventories (MCDI; Fenson, 2007). In general they found a steep increase on all language measures with increasing age. Girls were found to be significantly earlier than boys on word production, word comprehension and complexity. Additionally they found large variability between children of both genders on the different language-measures from the MCDI.

Pre-linguistic communication

The pre-linguistic phase refers to the period before the child starts producing words. It typically ranges from the child are newborn, up until the development of first words. During this period the newborn uses eye-contact in interaction with his/her caregiver as one of the first forms of early communication. As the newborn grows, so does the repertoire of non-verbal communication, including social-smiling and later proto-conversations (early turn-taking with smiling, body movement and making sounds). This later develops into communicative intentions, which includes use of gestures as an addition to the repertoire (e.g. pointing towards objects). By 8 months many children understand as many as 20 words. The

child has now started to understand segments in the caregiver's continuous speech, and associates these sounds with objects (Owens, 2012).

Both the parents and the child are active partners in early communication, with caregivers responding to the child's signals. The child's personality, emotions, preferences and interests for specific objects affects the development of language. Joint attention plays a crucial role during this period of language development as it sets the premises for communication and language learning thru everyday routines and play (Owens, 2012).

Early words and phrases

After the first year of life early words develop gradually, mixed with incomprehensible strings of sounds as the child experiments with new words. This period evolves as the child starts experimenting with walking. Vocabulary slowly grows, reflecting the world of the child. Context plays an important part in early word learning, as one study shows a strong connection between children's first words and maternal use of the same words (Harris, Barrett, Jones, & Brookes, 1988). However, this connection was much weaker for word learning beyond the initial first words (Barrett, Harris, & Chasin, 1991). By 18 months the child will have a vocabulary of about 50 words. The initial 100 word acquisition might progress slowly, but this period is often followed by a vocabulary spurt. At the same time as the children's vocabulary grows they start combining words. Kristoffersen et al. (2012) found that by the age of 19 months 50% of their Norwegian sample had begun to combine words, and by 22 months 75 % of their sample combined words. By the time a child is 24 months he or she can engage in conversations, as well as imitate and maintain focus on some topics (Owens, 2012). The child has started to use pronouns to refer to objects, but this is often done without identifying the object (e.g. "my thing" instead of "my car").

Pre-school language abilities

As the child reaches three years, the child's gross- and fine- motor skills develop further. The child now walks effortlessly on flat surfaces and experimenting with running, while using fine motor skills to dismantle toys. The child also uses language in more ways, and uses it to negotiate in play with other children. Usually children have an expressive vocabulary of about 900 to 1000 words at three years and many are seemingly chattering non-stop. As the child turns four years, motor skills continue to develop, the child can balance on one foot and some

might even be experimenting with writing letters. Role-play with other children becomes increasingly frequent during this period, and vocabulary has usually grown to about 1500 to 1600 words (Owens, 2012).

2.1.3 Relevant language functions

The following includes an introduction to relevant language factors/ functions discussed later on in the text.

Mean length of utterance

Mean length of utterance can be defined as the child's average utterance, either measured in morphemes or number of words. According to Owens (2012) the mean length of utterance (up to an MLU of 4.0) only acts as a moderately good predictor for the complexity of a child's language abilities. However, it has been proposed that MLU is a stable and reliable measure of language competence between ages 3 to 10 years (Rice et al., 2010; Rice, Redmond, & Hoffman, 2006). The use of MLU has also been recommended as a benchmark for measuring language intervention outcomes for children with autism spectrum disorders (ASD) (Tager-Flusberg et al., 2009).

Plurals

Development of plurals often follows a learning curve, where the child starts marking plurals with "more house", and later moves on to add the suffix –s to some words (e.g. houses). The child starts by adding the correct suffix to high frequency words the mother uses (Owens, 2012). Kristoffersen et al. (2012) found that 50 % of their sample mastered plurals (Norwegian suffix –er) by 23 months, and 75 % by 27 months.

Pronouns

Pronoun case marking errors in children two to four years old are one of the most typical language errors in English speaking children. They typically make statements such as: "me do it", or the more common "he wants milk" (instead of "I want milk") (Owens, 2012).

Past tense

Typically, when children have learned regular past tense rules (-ed) they start to overgeneralize (e.g. eated). Some children have an extended period of time where they use both correct past tense and overgeneralize. Over time the correct form gets established (Owens, 2012). In their study, Kristoffersen et al. (2012) found that 50 % mastered past tense (Norwegian suffixes –a,-et –te,-de) by 24 months, and 75 % mastered it by 28 months.

Narratives

The first emergences of narratives begin at age two. Narratives can be defined as self-generated stories that include sharing of relevant information, repairing broken dialogs and assuming the perspective of the listener. It is also essentially a monologue, where the language sets the context (i.e. it is not a context that is being shared, but language is used to explain the narrative to the other person). Narratives include two different approaches; centering and chaining. Centering is linking similar elements to form a story, while chaining consists of sequences of events that share similarities and lead from one to the other. At age two most children use centering of narratives, which focuses on the child's context and has a vague plot. By the age of three the child increasingly uses chaining (Owens, 2012). According to Botting (2002) constructing narratives are one of the more complex language functions, relying on both linguistic and pragmatic competence in the child. She also proposes that it is a reliable and good measure for distinguishing children with communication difficulties from typically developing children.

2.2 Autism spectrum disorders (ASD)

ASD is a neurodevelopmental disorder that affects social interaction, communication and includes stereotypic and restricted interests. In both the diagnosis manuals ICD-10 and DSM-IV-TR the term pervasive developmental disorder (PDD) is used as an overarching diagnostic category for all the subgroups, but the term autism spectrum disorders (ASD) is interchangeable with PDD.

There are some differences between DSM-IV-TR and ICD-10 in how they define sub-groups of ASD, but research has shown that there is high diagnostic consensus between the two diagnostic systems (Sponheim, 1996). The ABC-study, which the current study builds upon, uses the DSM-IV-TR diagnostic category of Autistic disorder, pervasive developmental disorder-not otherwise specified (PDD-NOS) and Asperger syndrome. The various sub-diagnoses within ASD refer to whether the child meets criteria for any of the twelve core symptoms specified under Autistic disorder, but each require a different number of criteria. The most significant difference between subgroups is between Autistic disorder and Asperger's syndrome, were the latter requires normal development of language (i.e. reaching language milestones within norms) and non-impaired cognitive abilities.

The current criteria for a 299.00 autism diagnosis in DSM-IV-TR are the following:

A. A total of six (or more) symptoms from the areas 1-social impairment, 2-impaired communication, and 3-repetitive behavior/interests, with at least two from (1), and at least one each from (2) and (3):

(1) Qualitative impairment in social interaction, as manifested by at least two of the following: (a) Marked impairment in the use of multiple nonverbal behaviors, such as eye-to-eye gaze, facial expression, body postures, and gestures to regulate social interaction. (b) Failure to develop peer relationships appropriate to developmental level. (c) A lack of spontaneous seeking to share enjoyment, interests, or achievements with other people (e.g., by a lack of showing, bringing, or pointing out objects of interest). (d) Lack of social or emotional reciprocity.

(2) Qualitative impairments in communication, as manifested by at least one of the following: (a) Delay in, or total lack of, the development of spoken language (not accompanied by an attempt to compensate through alternative modes of communication such as gesture or mime).

(b) In individuals with adequate speech, marked impairment in the ability to initiate or sustain a conversation with others. (c) Stereotyped and repetitive use of language or idiosyncratic language. (d) Lack of varied, spontaneous make-believe play or social imitative play appropriate to developmental level.

(3) Restricted, repetitive, and stereotyped patterns of behavior, interests, and activities as manifested by at least one of the following: (a) Encompassing preoccupation with one or more stereotyped and restricted patterns of interest that is abnormal either in intensity or focus. (b) Apparently inflexible adherence to specific, nonfunctional routines or rituals. (c) Stereotyped and repetitive motor mannerisms (e.g., hand or finger flapping or twisting or complex whole-body movements). (d) Persistent preoccupation with parts of objects.

B. Delays or abnormal functioning in at least one of the following areas, with onset prior to age 3 years: (1) social interaction, (2) language as used in social communication, or (3) symbolic or imaginative play.

C. The disturbance is not better accounted for by Rett's disorder or childhood disintegrative disorder (American Psychiatric Association, 2000).

The other sub-groups within the PDD group are defined through having fewer symptoms on any of the three core domains or in total number or showing later debut in symptom presentation (i.e. later than 36 months). However it is now commonly viewed that differentiating the various sub-diagnoses within the PDD-group is difficult. Research has suggested that there is high agreement with regards to whether the individual is within the autism spectrum or not, but differentiating between the different sub-groups is unreliable and influenced to a large degree by where the individual is assessed (Lord & Jones, 2012). This is the key reason for the changes suggested in DSM-5 which moves the field in the direction of using Autism Spectrum disorder (ASD) as the only diagnostic category. There will be specifiers for severity and function in those of diagnosed. The specifiers will indicate some of the differences currently found between subgroups (Lord & Jones, 2012). The DSM-5 is scheduled for release by mid of May 2013.

There have been many reports of gender differences in ASD. In a literature review by Mandy et al. (2012) they found ASD to be more common in boys than in girls, ranging from four to one in the whole spectrum, to eight to one in higher functioning samples (those without

cognitive difficulties). They also found that girls in the ASD spectrum on average had lower cognitive abilities than boys.

2.2.1 ASD and language difficulties

It has been suggested that language abilities is the characteristic that varies the most among people with ASD (Weismer, Lord, & Esler, 2010). It ranges from children with structural language (syntax and lexical) within normal limits and whose primary difficulties seems to be in the use of language in context and in communicating (semantics and pragmatics), to children who do not develop a vocabulary of any size, and have no functional language at all during early years (Volden et al., 2011).

Tager-Flusberg, Paul, and Lord (2005) showed in their literature review that the large variation in spoken language abilities were related to the child´s age and cognitive functioning. The language was appropriate for each child's mental age, and in general children with ASD who could talk showed unimpaired articulation, morphology and sentence structure. Additionally they found word use to be somewhat idiosyncratic, pedantic or overly concrete (this is often associated with pragmatics). It was also found that pronoun reversal was more common for young children with ASD compared to children with a general learning disability. Interestingly, pronoun errors are also present in typically developing (TD) children. While some TD children are having problems using their name instead of "I", most were related to the use of "you" instead of "me" (Evans & Demuth, 2012). Oshima-Takane, Goodz, and Derevensky (1996) found that second born children had greater abilities for producing and comprehending personal pronouns, compared to first born children. However the second born did not have better general language abilities. They hypothesize that second born children are more exposed to personal pronouns; because they hear the parents address their siblings with the same pronouns used towards them. Additionally, it has been assumed that pronoun acquisition is related to cognitive development, especially the child's understanding of itself and others (theory of mind). It has also been hypothesized that many children use pronouns before they understand the meaning, e.g. the child can for a period in development understand "you" as a name (Evans & Demuth, 2012).

Additionally, Mayo, Chlebowski, Fein, and Eigsti (2013) found that producing words before 24 months predicted better cognitive abilities and adaptive skills for children with ASD. They suggest that the current understanding of predicting better outcome for children with ASD

that develop useful language before 5 years should be updated. It is worth noting that even though producing words before 24 months predicted better outcome, the ASD group was still severely delayed. Maljaars, Noens, Scholte, and van Berckelaer-Onnes (2012) compared a low functioning ASD group with children with cognitive difficulties. Their study showed that lower receptive and expressive scores were found in the ASD group compared to the other group. Additionally there were bigger discrepancies between non-verbal IQ (NVIQ) and both expressive and receptive language scores for the ASD group. The ASD group showed more uneven language abilities than the group with cognitive difficulties.

There have been several studies investigating specific language components in ASD. According to Norbury and Bishop (2003) narratives is a known difficulty for ASD, and even for children with SLI. While others have hypothesized that narratives are related to pragmatic competence, they found core language ability to be more strongly related. Tager-Flusberg and Joseph (2003) investigated children with ASD on, among other tests, a past tense tasks. They found that a portion of the children with ASD omitted the morphological marking on verbs (e.g. saying walk for walked). Lastly, Eigsti, de Marchena, Schuh, and Kelley (2011) found that several studies have shown vocabulary skills to be predictive for later language development in ASD. However other studies have shown that NVIQ is the most significant predictor associated with impaired language functioning (Maljaars et al., 2012; Weismer et al., 2010).

2.3 Conceptualizing language impairment (LI)

Language impairment (LI) is associated with a wide array of different developmental disorders. LI can exist as the child's only difficulty, or coexist with others, like autism or motor difficulties. There is disagreement in the literature concerning what is the most appropriate term, and what defines language impairment as specific (Bishop, Clark, Conti-Ramsden, Norbury, & Snowling, 2012; Conti-Ramsden, 2009). Traditionally, specific language impairment (SLI) is defined as children with primary difficulties in the language domain and with cognitive abilities within normal limits (non-verbal IQ >85). Distinguishing children with a general learning impairment from children with SLI based on a non-verbal IQ measure seems possible as one approach, but a too strict understanding of this has been found to be problematic. Tomblin et al. (1997) found that some children with NVIQ below 85 showed a clinically indistinguishable language pattern from children with higher IQ levels.

Botting (2005) reviewed several studies on SLI, and found NVIQ to fluctuate or drop within a couple of years in clinical populations of children diagnosed with SLI. Little is known of the direction of the relationship between language and NVIQ. It has been hypothesized that the language impairment limits development of NVIQ, as well as the other way around.

Typically, children with language difficulties are classified in DSM-IV-TR and ICD-10 diagnostic manuals as having phonological problems and/or expressive problems with or without receptive difficulties.

The current diagnostic criteria for expressive language disorder (315.31) in DSM-IV-TR are:
(a) The scores obtained from standardized individually administered measures of expressive language development are substantially below those obtained from standardized measures of both nonverbal intellectual capacity and receptive language development. The disturbance may be manifest clinically by symptoms that include having a markedly limited vocabulary, making errors in tense, or having difficulty recalling words or producing sentences with developmentally appropriate length or complexity.
(b) The difficulties with expressive language interfere with academic or occupational achievement or with social communication.
(c) Criteria are not met for mixed receptive-expressive language disorder or a pervasive developmental disorder.
(d) If mental retardation, a speech-motor or sensory deficit, or environmental deprivation is present, the language difficulties are in excess of those usually associated with these problems (American Psychiatric Association, 2000).

Additionally the current diagnostic criteria for mixed receptive-expressive language disorder (315.32) in DSM-IV-TR are:
(a) The scores obtained from a battery of standardized individually administered measures of both receptive and expressive language development are substantially below those obtained from standardized measures of nonverbal intellectual capacity. Symptoms include those for expressive language disorder as well as difficulty understanding words, sentences, or specific types of words, such as spatial terms.
(b) The difficulties with receptive and expressive language significantly interfere with academic or occupational achievement or with social communication.
(c) Criteria are not met for a pervasive developmental disorder

(d) If mental retardation, a speech-motor or sensory deficit, or environmental deprivation is present, the language difficulties are in excess of those usually associated with these problems (American Psychiatric Association, 2000).

Some research has shown that the current classification conceals additional impaired language functions. According to Conti-Ramsden, Crutchley, and Botting (1997) this includes several groups, including expressive phonological impairment with comprehension relatively unimpaired, mixed receptive-expressive phonological and syntactic impairment and lexical-syntactic problems with word-finding difficulties that includes immature syntax. They also described one group with semantic and pragmatic impairments with unimpaired phonology and syntax. Additionally it has been found that children with SLI are at risk for other comorbid conditions such as developing reading disabilities and psychosocial difficulties (Snowling, Bishop, & Stothard, 2003; Snowling, Bishop, Stothard, Chipchase, & Kaplan, 2006) and comorbid motor difficulties are not uncommon (Wang, Lekhal, Aarø, & Schjølberg, 2012; Webster & Shevell, 2004). Lastly, early language impairment has also been associated with risk for academic failure in school, and some studies have shown that early intervention did strengthen the child's reading development (Fricke, Bowyer-Crane, Haley, Hulme, & Snowling, 2013).

There have also been several studies investigating specific language functions. Conti-Ramsden and Hesketh (2003) investigated acquisition of plurals in children with SLI. They found it had low strength in predicting SLI, and found that children with SLI were almost as good as typically developing children. Their sample was between the age 52 and 70 months, a bit older than the participants in the current study. Another study of German children with SLI between the ages of 3 and 6 years found that they performed in line with MLU matched controls (Kauschke, Kurth, & Domahs, 2011). Similar to the Conti-Ramsden and Hesketh (2003) study they found that when the children made errors they seemed to overuse the –s suffix. Delayed vocabulary is also highly associated with LI (Hick, Joseph, Conti-Ramsden, Serratrice, & Faragher, 2002). Rice, Wexler, Marquis, and Hershberger (2000) investigated children with SLI acquisition of past-tense. They found that children with LI have a prolonged period of acquisition for regular past tense, and when making errors they often omitted the morphological marking on the stem of a verb (e.g. saying walk for walked). This is similar to what Tager-Flusberg and Joseph (2003) found for children with ASD.

Rice (2013) examined language development compared language growth curves in children with and without SLI using growth curve models. She found that children with SLI were delayed, but had the same growth curve model, across many different language functions as children without SLI. The growth models were different depending on the language function assessed, but showed a similar pattern; the SLI group had delayed onset relative to age expectations, but when acquisition had started it seemed to follow the same growth pattern of language acquisition as children without SLI. However, while the growth pattern was the same it seemed to level off at a lower threshold than the children without SLI, meaning their language levels where bellow age expectations. She found a general delay in onset of about two years for all language functions assessed. In sum the author argues that this shows a likely surprising biological drive strength in the language acquisition when the growth first has started, similar both in for children with/without SLI (Rice, 2013).

There seems to be few studies that have investigated possible gender differences in LI. One study by Tomblin et al. (1997) investigated the prevalence of SLI in pre-school children. Of 216 children with SLI, he found 59 % to be male, and 41 % to be female. This is different from previous findings of a ratio of 2:1 for boys. However, there does not seem to be any descriptions of potential differences between the genders in language functioning in the literature.

In summary, children with language difficulties are a highly heterogeneous group, and therefore language impairment (LI) could be seen as a more appropriate term to describe children with various language difficulties. In the current study the term LI will be used to describe the language impaired participants who did not have ASD, severe intellectual disorder or any syndrome comorbid diagnosis that seem to explain their language impairment. The research literature sometimes use the label differently, however in the review of the literature the use by the various researchers will be maintained.

2.4 Overlap in language between children with LI or ASD

Several researchers have described an overlap in language problems seen in the LI and ASD groups. Traditionally, difficulties with pragmatic language are seen as the hallmark for children with ASD. Pragmatic difficulties are often used as a way of differentiating ASD from SLI. However, a wide array of literature describes a substantially overlap between SLI and ASD (Bishop, 2003; Bishop & Norbury, 2002; Conti-Ramsden, Simkin, & Botting, 2006). Many researchers view SLI and ASD as being on the same continuum, where pragmatic language difficulties links the impairments (Bishop, 2003). Additionally, it has even been proposed that a separate subcategory of pragmatic language disorder exists outside the autism spectrum (Bishop, 2003; Botting & Conti-Ramsden, 1999). The development of DSM-5 has proposed a new diagnostic category within the language impairment groups called social communication disorder, roughly corresponding to the semantic-pragmatic language impairment proposed by Bishop (2003), but without the ASD symptom pattern (Lord & Jones, 2012).

Little is known about how the proposed overlap in language problems between ASD and LI groups is manifested in young children. In a literature review by Williams, Botting, and Boucher (2008) they found that by school age the language abilities in the ASD group and the SLI group were significantly different. However, some overlaps were found, especially in preschool. Rapin and Dunn (2003), on the other hand, found, indications that the SLI group had more "pure" expressive language difficulties than the ASD group in their study of preschool children. They also found that the ASD group had difficulties with semantics and pragmatics.

Two studies by Kjelgaard and Tager-Flusberg (2001) and Tager-Flusberg and Joseph (2003) compared autism symptomology and language abilities in ASD and SLI. They describe a possible theoretical framework for understanding the proposed overlap between the diagnoses. According to the authors ASD and SLI manifest themselves as two separate diagnoses, but they can co-occur and act as a double-hit in autism that they called ALI. This double-hit theory has been hypothesized to lead to increased levels of impairment for the individual, and is related to higher scores on autism symptomology. However, Loucas et al. (2008) did not find that having ALI lead to greater autism symptomology, but did find greater

impairment of language function. Additionally, they found that the SLI group had several positive scores on the Autism Diagnostic Interview - Revised (ADI-R; Rutter, Le Couteur, & Lord, 2003), but they still had lower scores on ADI-R than the ASD group. This might reflect some similarities between the groups. When comparing language abilities they found the ASD group had equally impaired receptive and expressive language, while the SLI group had stronger receptive language than expressive.

2.5 Early language characteristics in ASD and LI

In general there seems to be limited information in the literature on early language characteristics in children with ASD as young as three years (See Charman, Drew, Baird, & Baird, 2003; Luyster, Lopez, & Lord, 2007; Weismer et al., 2010). In these studies the majority of children with ASD exhibited significant delays in each of the areas assessed, including vocabulary comprehension and production, nonverbal communication skills, functional object use and play skills. In addition, Weismer et al. (2010) reviewed findings from other studies (See Eaves & Ho, 2004; Mitchell et al., 2006; Paul, Chawarska, Cicchetti, & Volkmar, 2008). In these studies there was substantial variation in the development of spoken language and presence of atypical language characteristics within the ASD group. Mental age might at least explain part of this variation. Interestingly, Leyfer, Tager-Flusberg, Dowd, Tomblin, and Folstein (2008) found that children with SLI also had scores on some items of atypical language characteristics often found to be associated with autism. In their sample they found that 41% of the SLI group met autism or autism spectrum cut-offs for social or communication domains on the ADI-R.

In summary, the language difficulties found in both ASD and SLI seems to be understood as an indication of a developmental overlap between the two groups. Whether this is a sign of a shared etiology or just shared clinical features remains unknown.

2.6 Social difficulties in ASD and LI

Social difficulties are one of the core symptom areas in ASD. Examples range from difficulties with social interaction in general, to difficulties in developing peer relations. According to Stanton-Chapman, Justice, Skibbe, and Grant (2007) positive peer interactions have been characterized as the child's abilities to engage and initiate conversations, communicate intentions clearly, address all children when joining a group and making adjustments in communications with regard to the listeners. These functions rely on language and pragmatic abilities alike. While the social difficulties for ASD are well known, the social difficulties related to LI are less documented. Given the language difficulties in LI it is not surprisingly that this limits the child's abilities to interact with peers. Fujiki, Brinton, Isaacson, and Summers (2001) found significant differences in peer interaction, and withdrawal behavior in a sample of school children with SLI. The children were found to spend less time interacting with their peers than typically developing (TD) children. In a study looking at pre-school children, Stanton-Chapman et al. (2007) found that children with SLI scored significantly different on internalizing behaviors (e.g. anxiety/depression and withdrawal) and social skills (e.g. cooperation, responsibility) compared to TD children, but there were no differences on externalizing behaviors (e.g. attention problems and aggressive behaviors). Another study by Hart, Fujiki, Brinton, and Hart (2004) found that level of language impairment was associated with social behavior. Children with less language impairment showed higher levels of social behavior, however, severity of language impairment was not associated with withdrawn behaviors.

Long term outcome of language impairment has also been investigated. St Clair, Pickles, Durkin, and Conti-Ramsden (2011) followed a sample of children with SLI at 7, 8, 11 and 16 years. They found that there was a decrease in behavioral and emotional problems from childhood into adolescence, although emotional problems were still present. However, they also found an increase in social problems for the adolescent. Another study by Johnson, Beitchman, and Brownlie (2010) followed children who had been identified with SLI at the age of 5 years, and were seen again at 12, 19 and 25 years. This study showed that adults at age 25 with a history of SLI had poorer outcomes in communication, educational attainment and occupational status. However, they did find similarities with the control group. Perhaps the most notable was the quality of life self-evaluation variables, where they found no differences between those with and without SLI.

Theoretical models concerning the relationship between social difficulties and language impairment are few. However, it has been hypothesized that social problems stem directly from the language problems children experience when they engage in social interaction. This model is called the social adaption model (Redmond & Rice, 1998) which hypothesize that children with LI withdraw from social interactions because of their language problems, thereby limiting their experience with peer interactions. Bishop (1997) speculated that an underlying cognitive deficit is the reason for the social difficulties as exemplified in limitations in working memory and low capacities in speed of processing. The child might have difficulties with processing the amount of linguistic information and therefore showing social difficulties. It might be possible that the same cognitive difficulties contribute to the social difficulties in ASD. Lastly it has also been hypothesized that children with LI have specific social or emotional deficits that contribute to their social difficulties (Hart et al., 2004).

Orsmond, Shattuck, Cooper, Sterzing, and Anderson (2013) investigated social participation in young adults with ASD (age 21 – 25). They found that young adults with ASD were more likely to be male, have higher rates of conversational impairment and likely to live under supervision than peers with other types of difficulties (cognitive difficulties, emotional disturbance and learning disabilities). With regard to peer-relationships they were significantly less likely to see friends, get called by friends, get invited to activities and were found to be socially isolated. Russell et al. (2012) found that children identified at an early age had worse outcome than children who showed diffuse autism symptomology in pre-school, and were identified later. They hypothesized that the sample that is identified at an earlier age is more severely impaired, and therefore have poorer long-term outcome. This could also be seen in connection to average age of diagnosis. Mandell, Novak, and Zubritsky (2005) found, via questionnaires sent to parents, that the average age of diagnosis was 3.1 years for children with autistic disorder, 3.9 years for PDD-NOS, and 7.2 years for Asperger's disorder. The results from this study shows there are big differences in age of diagnosis between the current subgroups in ASD. This again could be related to level of impairment, as children with Asperger's disorder have better language and cognitive abilities than children with autistic disorder.

3 Method

The current study used data collected in the Autism Birth Cohort Study (ABC-study). The ABC-study aims to investigate causes of autism and examine how ASD develops in children (Stoltenberg et al., 2010). The ABC-study is a sub-study of the Norwegian Mother and Child Cohort Study (MoBa; Magnus et al., 2006). MoBa is a pregnancy cohort that includes 90 700 mothers, 72 100 fathers and 108 500 children. MoBa recruited participants from 1999 to 2008 from various birth clinics in Norway. Mothers in the MoBa cohort answer questionnaires during pregnancy, and at given intervals after the child's birth (the questionnaires so far are: 15th, 22nd and 30th week of pregnancy, 6-, 18- and 36-months after birth and at 5-, 7- and 8 years of age). The clinical assessment in the ABC study is cross-sectional, inviting children once for a developmental and diagnostic assessment. In the current study only clinical data from the ABC-study was used, not the longitudinal data from MoBa. However, understanding MoBa is important to comprehend how the participants were recruited to the ABC-study.

3.1 Participants

Participants in the ABC-study are recruited using four different methods: (1) screening via the 36-months questionnaire, (2) professional referrals from health practitioners, (3) self-referrals from parents and (4) linkage with the Norwegian Patient Register (NPR). As the study progressed it also included screening via questionnaires at 5- and 7 years of age. In addition, the study invited age-matched controls that were randomly selected from the MoBa cohort and matched to potential ASD-cases by date of birth (Stoltenberg et al., 2010).

Screening from the 36-months questionnaire included using the Social Communication Questionnaire (SCQ; Rutter, Bailey, & Lord, 2003). SCQ consists of 40 questions from the first edition of the Autism Diagnostic Interview-Revised (ADI-R), and has been validated for screening for autism in clinical populations. Seven questions were excluded in the screening algorithm due to requiring the child to combine words or using sentences. The number of questions used in the algorithm was therefore reduced to 33. Scoring was either 0 or 1, where 0 indicated normal development (Unpublished ABC-study protocol, 2010).

In 2006, at the beginning of the study, children screened positive if they meet one or more of the following criteria: (1) total SCQ-33 score $\geq$ 12, (2) a full score (9 out of 9 points) on the

repetitive behavior domain of the SCQ, (3) parents report language delay AND the child has been referred to a specialist for it, (4) parents report autism/autistic traits OR the child has been referred to a specialist for it, (5) parents report that the child shows very little interest in playing with other children. In addition a control group of randomly selected, age matched children were drawn.

To increase specificity for ASD, a slightly revised version of the screening algorithm was used from February 2007. An additional item was added to the list, namely (6) parents report that others (family, daycare staff, well-baby nurse) have expressed worry for the child's development. The children now screened positive if they met criterion 4 (autism/autistic traits) or met one or more of the "old" criteria in addition to criterion 6.

Participants in the current study were selected on the basis of participation in the ABC-study. For investigating early language characteristics the age range was determined to be from 3.0 to 4.11 year as these are the youngest children in the ABC-sample. The diagnoses of the participants in the current study were established at the ABC-clinic. The groups included in the current study are the diagnostic groups ASD and LI, as well as typically developing (TD) children. The ASD group included the following diagnoses: autistic disorder, Asperger's syndrome, and pervasive developmental disorder – not otherwise specified (PDD-NOS). The LI group included: expressive language disorder and mixed receptive-expressive language disorder. The participants who had a primary diagnosis of phonological language disorder (N=42) were excluded due to problems in determining whether their problems were purely articulatory or of a more severe phonological form. The TD group consisted of children who did not receive any diagnosis at the ABC-clinic and did not show any clinically significant problems.

3.2 Instruments

The ABC-study includes many different instruments for assessing child development. In the current study, only a sub-set of cognitive tests, interviews and questionnaires have been used. For a complete overview of the instruments in the ABC-study see Stoltenberg et al. (2010). The following section will include a presentation of the instruments used in the current study.

3.2.1 Cognitive tests

Stanford-Binet Intelligence Scale 5th ed. (SB5; Roid, 2003) is administered individually to the participant. It consists of verbal and non-verbal "routing" tests (a short version of the full battery that gives an abbreviated IQ score), six nonverbal subtests and five verbal subtests. From 2005-2008, and from June 2010 onwards, the full version of the SB5 was administered in the ABC-study. From January 2009 to July 2010, only the following subscales were included: nonverbal fluid reasoning (nonverbal routing), verbal knowledge (verbal routing), verbal fluid reasoning, verbal working memory and non-verbal working memory. From these scales an abbreviated IQ can be calculated. Since the ABC-study added an additional non-verbal scale, an abbreviated nonverbal IQ can be calculated. The SB5 is standardized for use from ages 2 to 85 years, but there has been reported certain limitations when it is used with very young children with developmental delay (Unpublished ABC-study protocol, 2010). SB5 is translated for use in the ABC-study, and uses American norms as SB5 is not traditionally used in Norway.

Mullen Scales of Early Learning (Mullen; Mullen, 1995) is used as a cognitive measure for children that are unable to complete SB5. Mullen is designed for infants and preschoolers from birth through 68 months of age. It consists of tests that measure gross motor, visual reception, fine motor, receptive language, and expressive language. Overall results are referred to as an early learning composite score. A study by Wetherby et al. (2004) found that many young children with ASD scored the lowest score possible on the early learning composite in Mullen.

3.2.2 Interviews

Autism Diagnostic Interview-Revised (ADI-R; Rutter, Le Couteur, et al., 2003) is a semi-structured parental interview containing 90 questions, that covers a range of social, language, and repetitive behaviors, as well as other features common in autism but not part of the diagnostic criteria. The subjects are rated for both "now" and "ever" scores, where "now" refers to the current situation, and "ever" refers to the most severe period in time. Additionally, some "ever" scores are substituted with a "4 – 5 most abnormal" score. This is an age period where it is hypothesized that ASD symptomology is at its most severe. The 4-5 scores function as a cross-sectional score for comparing symptomology for children interviewed at different ages. The scores have different criteria and wording, but there is a

general “theme” in each rating scale: “0” specifies absences of the particular behavior in question, “1” specifies presence of possible autistic-like behavior (i.e. not definite), “2” indicates definite presence of the behavior and moderate severity, “3” indicates definite presence of the behavior and high-severity with impact for the child and/or family. In the current study only a selection of language related questions was used as a measure of presence or absence of abnormal language behavior, although the whole ADI-R was administered at the time of the visit.

Vineland Adaptive Behavior Scales – communication subdomain (VABS; Sparrow, Balla, & Cicchetti, 1984) is a semi-structured parental interview that measures the child’s functional communication ability in daily living. The questions are open-ended, and gather descriptions and examples of the child’s daily communication behaviors, which in turn are coded by the interviewer according to the instrument manual. The scores range from “0”, indicating absence of daily use, “1”, indicating the child uses the skill sometimes, to ”2”, which indicate everyday use. The VABS is translated from the American version for use in the ABC-study and uses American age norms. In the current study only the expressive and receptive scores from the communication domain will be used.

3.2.3 Questionnaires

The Child Development Inventory – expressive language subscale (CDI; Ireton, 1992) is a 50-item parental-report questionnaire. It is used as a measure of the child’s expressive language abilities. It is comprised of yes and no answers to questions regarding the child’s expressive language abilities. Chaffee, Cunningham, Secord-Gilbert, Elbard, and Richards (1990) found that the Minnesota Child Development Inventory (the predecessor to the CDI) was an effective screening instrument in language-problem based populations (i.e. it identified correctly children with language difficulties), but had a tendency to produce false-positives (i.e. reported difficulties for children without language disorders). The Norwegian version of the questionnaire is included in the appendix E.

The Preschool Play Behavior Scale (PPBS; Coplan & Rubin, 1998) is a 21-item questionnaire regarding the child’s behavior during play. In the current study it is used as a measure of how the child uses language socially during play, and how the child is engaging with other children. It contains a 5-point Likert response-scale, ranging from “never” to “very often”. The questions are clustered into five different factors: reticent behavior, solitary-passive

behavior, solitary-active behavior, social play and rough-play. The original article (Coplan & Rubin, 1998) found that the PPBS was a good alternative to observational descriptions of children's play behaviors. In the current study, only two of the factors were used: solitary-passive and social play. The PPBS was included in the parental questionnaires from 2006 to 2009.

3.3 Procedures

The following section includes descriptions of the assessment procedure from the ABC-clinic, were the participants were assessed. Additionally, this section includes how the various measures in the current study were constructed. Lastly the statistical analysis will be presented.

3.3.1 Assessment at the ABC-clinic

The screened participants from the 36 months questionnaire were invited to take part in a one day assessment at the ABC-clinic located in an office building at Lovisenberg Hospital in Oslo. The clinic was operated by Nic Waals Institute in collaboration with the Norwegian Institute of Public Health (NIPH). All children were assessed by a team consisting of a psychologist, a child psychiatrist and research assistants.

Three weeks prior to the scheduled appointment at the ABC clinic the parents received a pre-assessment form which contained a set of questionnaires (including the PPBS and the CDI). The parents also receive a second pre-assessment form, which they were asked to deliver to the daycare.

A day at the ABC-clinic was normally split into two parts. Before lunch the child was assessed with a test battery including psychometric testing with SB5 or Mullen, fine and gross motor tests, a semi-structured play-based observational test and a mother and child interaction sequence. The tests were administered by a clinical psychologist with a parent present in the room. After lunch, the day included taking a blood sample and various interviews of the mother while the child was able to take a break. For the mothers this included talking with a child psychiatrist about child and family medical history and two interviews, the ADI-R and VABS, regarding the child's development. The interviews were administered by a trained research assistant certified in the instrument used. Most of the sequences were video-taped for

quality-assurance, and future research. Some variations of the arrangement of the day were done with regard to the child's needs and efficiency of the clinic, but in general the day included the same components in the order described above.

At the end of the day the assessment team led by the child psychiatrist or a specialist in clinical psychology reviewed all available information before concluding with a possible diagnosis in accordance with the DSM-IV-TR manual. Before leaving the clinic the parents received feedback of the results. The child was not present during this time. In addition to the feedback at the end of day, the parents received a written report by mail. The parents could use this report when contacting local support institutions for further assessments and/or interventions for the child. The parents also had the opportunity of calling the clinic to talk to one of the clinicians from the assessment team if they had questions, both prior to the assessment as well as after.

The participants received reimbursement for travel, accommodation and food expenses. No other monetary benefits were given for participating.

3.3.2 Construction of measures

The following includes a presentation of how the various measures used in the study were constructed. The construction of the presented variables was done specifically for the current study.

Cognitive composite measure

To get a full-scale IQ measure (FSIQ) on as many subjects as possible, a mental age/cognitive composite measure was made by combining different standardized scales from SB5 and Mullen into one variable. See table 1 for an overview of the number of participants for a given measure. A total of four different measures were used. The two SB5 measures (full-scale IQ and abbreviated IQ), as well as the Mullen Early Learning Composite standard scores, are calculated from the respective test-software. The Mullen converted T-score is derived from the mean value T-scores of two verbal and two non-verbal subtests. This score was then converted from a T-score to a standard score, with a mean of 100, using a conversion table (Psychometric Conversion Table, 2003).

When constructing the composite IQ measures a hierarchical approach was used to ensure that the best available measure, for any given participant, was used. The best available measure will depend on what instrument the participant completed at the ABC-clinic (i.e. SB5 or Mullen). The SB5 full-scale IQ is considered a more precise measure than the abbreviated IQ, due to the full-scale using additional subtests from SB5. The abbreviated IQ is based on the two routing subtests from the SB5. The two FSIQ scales from SB5 show a good level of internal consistency .831, as measured by Cronbach's alpha.

For participants completing the Mullen, the Early Learning Composite is considered a more precise measure than the Mullen converted T-score, as it uses more subtests and has standardized measures for the four subtests. The Mullen converted T-score is not a measure listed in the manual, but it is used as the best available measure if one is not able to get complete enough data for a given child. Both measures derived from Mullen show a good level of internal consistency .797, as measured by Cronbach's alpha.

In addition to the FSIQ measure, a separate composite measure was created for nonverbal IQ (NVIQ). The NVIQ score was constructed from three different measures (see table 1). The only manual-based measure for NVIQ is derived from the full-scale SB5. If the participant only completed the abbreviated SB5 with the additional nonverbal sub-test, an estimated NVIQ was constructed from the mean of the two scaled scores of these two nonverbal subtests (fluid reasoning and working memory). The score was then recoded with a mean of 100. Both the NVIQ scales from SB5 show a good level of internal consistency .868, as measured by Cronbach's alpha.

For participants tested with Mullen, a NVIQ was constructed from the mean T score of two subtests (fine motor and visual reception). The score was then converted from a T score to a standardized score, with a mean of 100, using a conversion table (Psychometric Conversion Table, 2003). The two scales show a good level of internal consistency .791, as measured by Cronbach's alpha

Table 1.
Instrument; Number of participants with FSIQ and NVIQ measures, across diagnostic groups

Instrument name	ASD group *(N=83)*	LI group *(N=161)*	TD group *(N=414)*
FSIQ instruments			
SB5 full-scale IQ	21	99	330
SB5 abbreviated IQ	39	60	84
Mullen Learning Composite	22	2	0
Mullen converted T-score	1	0	0
NVIQ instruments			
SB5 nonverbal	26	107	334
SB5 estimated NV	16	44	80
Mullen converted NV	41	9	0

The majority of the sample completed the SB5 full-scale or the SB5 abbreviated IQ test (N=633), while a small number completed the Mullen Learning Composite (N=24) and only one participant has a Mullen converted T-score. Most of the participants who completed the Mullen were in the ASD group. This is expected due to findings in the literature that young children with ASD are not able to complete a conventional cognitive test like the SB5, especially at an early age (Wetherby et al., 2004).

For the NVIQ composite measure, most participants had completed the SB5, and hence a nonverbal scale could be used. Many participants also had the SB5 estimated NVIQ, while fewest participants had the Mullen converted NV scale. This means that most of the participants had the most reliable measure available. However, most of the ASD group only had the Mullen converted NV score, which is considered the least reliable. The challenges of constructing these measures will be addressed in section 3.4.

Language development measures

For measuring language development across the various groups, the receptive and expressive age equivalent from Vineland was recoded into a DQ (development quotient) measure. This was done by dividing the age equivalent score with the participant's chronological age in months, and then multiplied with 100. The same was done with the age equivalent score from the CDI expressive scale.

For measuring developmental language milestones, two questions from the ADI-R were used: 1) age of first word and 2) age of first phrase. The mean age was calculated. When accurate month of achieving this skill was not available, the ADI-R manual states that parents should be asked to estimate if the development was "believed to be normal, months unknown", "believed to be delayed, months unknown" or "milestone not achieved". These categories will be presented.

Language component measures

Various language component measures were constructed from the CDI based on the underlying theme of the question. The participants were compared by age group (i.e. 3 and 4 years) due to developmental differences between these two ages. The items measuring a specific theme were summarized into a composite measure. Some of the scales were composed of items alluding to interval (i.e. uses 4 or more words, uses 8 or more words etc.), while others were at an ordinal level measuring presence of abnormalities. For testing if the items summarized were reasonable to consider as part of a coherent scale, Cronbach's alpha analysis was conducted for each composed measure. This was done on the complete sample (i.e. not split by age).

The length of utterance factor is made up of three questions: "talks in sentences that are at least four words long", "talks in longer sentences to express complete thoughts- at least six words long" and "talks in long, complex sentences, ten words or longer". This scale had an acceptable level of internal consistency .769.

The vocabulary cluster is made up of three questions: "uses at least 10 words", "has a vocabulary of 20 or more words" and "uses more than 50 words in everyday conversations". The scale showed an acceptable level of internal consistency .612

The grammar skills factor contains three questions: “uses plural words, the Norwegian equivalent of adding «s», for example, «girls», «cars»”, “talks in the past tense correctly, for example, says «I played with Billy.», «I did.», «We went...»” and “uses plurals correctly, for example, says «men», not «mans», «mice», not «mouses». This scale also had an acceptable level of internal consistency .726.

The pronoun factor contains four questions: “refers to his (her) things as «my» or «mine», “uses the word «you» in sentences”, “uses the words «me», «my» and «I» correctly” and “uses plural pronouns such as «we», «you», «they», «them» or «us» correctly”. The scale had an acceptable level of consistency .716.

Finally the narrative factor included two questions: “tells what action is going on in pictures - for example, «Kitty is eating»” and “retells short stories such as Little Red Riding Hood; tells what happens in correct order and how the story ends”. The scale has a low level of internal consistency .338.

Atypical language measure

For comparing the presence of atypical language often associated with ASD, across the three groups a set of language related questions from the ADI-R was used. Individual scores on specific questions were examined. The language cluster selected is derived from Tadevosyan-Leyfer et al. (2003), who did a factor analysis on all items from the ADI-R. Some items were not applicable because an earlier version of the ADI-R was used in their study.

Autism diagnostic symptoms

To compare autism diagnostic symptoms the domain algorithm scores from ADI-R was used. The algorithm is made up of scores from a number of questions from ADI-R, and reflects the core symptoms for ASD as determined in ICD-10 and DSM-IV-TR. The algorithm is split into three domains: qualitative abnormalities in reciprocal social interaction, qualitative abnormalities in communication, and restricted, repetitive and stereotyped patterns of behavior. The items from each algorithm domain questions were summarized into the domain score. Each domain also has a cut-off score that is meant to distinguish between potential children with ASD (above cut-off) and those without ASD (below cut-off) (Rutter, Le

Couteur, et al., 2003). The cut-off scores have been constructed statistically, by finding the optimal combination between specificity and sensitivity (Lord, Rutter, & Le Couteur, 1994).

Two different ADI-R algorithms are used in the current study (age 2:0 years to 3:11 years and 4:00 years or more). The algorithms reflect different developmental demands given the subject's age. The "4:00 years or more" algorithm has greater demands related to development of peer relationships, and therefore has three different items in the social domain compared to the algorithm for younger children. For a complete overview of both algorithms see the appendix D. Additionally, the participants were grouped into verbal and non-verbal participants depending on their score on item 30 in the ADI-R. A participants is considered verbal according to the ADI-R manual if the child uses utterances of three or more words daily, that includes a verb and is understandable for others (Rutter, Le Couteur, et al., 2003).

Lastly participants from all groups on the ADI-R were compared with above cut-off scores from both ADI-R and Autism Diagnostic Observation Schedule (ADOS; Lord, 2002) were identified. The ADOS is a semi-structured assessment instrument which evaluates the child's ability to engage in social interaction, communication, play and imaginative use of materials (Lord, 2002). It is often used in conjunction with the ADI-R. The ADOS has three different modules; the choice of module is related to the child's age and verbal abilities. Each module also has an algorithm with a cut-off score which indicates whether the child potentially has autism or autism spectrum disorders (with reference to the DSM-IV-TR and ICD-10 criteria). For the current study, the autism spectrum disorder cut-off from module 2 was chosen for verbal participants, and cut-off for autism spectrum disorder from module 1 was chosen for non-verbal participants. If the child was considered verbal or non-verbal was determined by item 30 in the ADI-R (as described above). There is no official cut-off score for repetitive behavior on ADOS, but for the current study it was determined that a mean value (3) of the scale should be used for both modules.

Play behavior measure

A measure of how the child engages in social behavior in pre-school was used. This measure consists of two clusters of questions from the PPBS scale (see table 2).

The clusters were “social engagement/interest in others play”, and “tendency to be solitary engaged rather than play with peers”. The questions were selected in accordance with the original paper from Coplan and Rubin (1998).

Table 2.
PPBS; questions included in the social play and solitary-passive factors

"Social play"	"Solitary - passive"
Plays «make-believe», but not with other children	Plays alone, building things with blocks and/or other toys
Engages in pretend play with other children	
Engages in active conversations with other children during play	Plays by himself/herself, examining an object or toy
Plays in groups with (not just beside) other children	Plays alone, exploring toys or objects, trying to figure out how they work
Talks to other children during play	Plays by himself/herself, drawing, painting pictures or doing puzzles
Engages in group play	

The overall theme of the “social play” factor is group play and interaction with others. It shows a good internal consistency of .897. The theme for the “solitary – passive” factor is engaging in playing alone and exploring objects. This factor shows an acceptable internal consistency of .760. The internal consistencies were measured by Cronbach’s alpha. There were low correlations between the two play behavior factors $r = .12$, p (two-tailed) $< .001$. This indicates that the two factors are measuring different aspect of play behavior that can coincide with each other but not necessarily.

3.3.3 Statistical analysis

All analyses were conducted using IBM SPSS Statistics for Windows, version 20.0 (IBM Corp, 2011). For investigation of group differences an analysis of variance (ANOVA) and analysis of covariance (ANCOVA) was used. An ANOVA is considered similar to a t-test, because we are comparing mean scores, but running several t-tests when comparing more than two groups inflates error rates. However, the ANOVA includes measures that avoid inflating the error rates. According to Field (2009) an ANOVA is basically the same as running a multiple regression analysis. An ANCOVA is in many ways the same as an ANOVA, but it makes it possible to control for variables that may influence the dependent variable. The variable controlled for is often called a covariate (Field, 2009).

For all measures the test of homogeneity of variance, measured by Levene's statistics, was significant, meaning that the variance within groups was different from each other. As a result Games-Howell post hoc test for ANOVA was used, but Field (2009) recommends using Bonferroni as a post hoc test for ANCOVA analysis. For the ANCOVA the homogeneity of regression slopes were investigated, and for most analysis in the current study this assumption was met. However, if the assumption was violated, an ANOVA was used instead. An effect size for between groups was calculated manually for all ANOVA analysis using the following equation: $\eta^2 = \frac{SS_{between}}{SS_{total}}$. Lastly the assumption of normality was not violated for most variables (for complete discussions see section 4.1 and table 3).

NVIQ was used as a control variable when running the ANCOVA analysis, for the following measures: language functions, language used in everyday living, autism symptomology, and play behavior. Since the groups in the current study are uneven and the child's mental age will influence the developmentally sensitive outcome measures, controlling for a possible cofounder (such as IQ) was considered needed. Since this study compares language measures we used non-verbal IQ as a confounder. Both FSIQ and VIQ are influenced by the difficulties with language, and therefore do not represent a good measures of cognitive functioning (Eigsti et al., 2011).

The atypical language measure was cross tabulated with diagnostic group, producing number of participants and percentages positive for each item. The same was done for the language component measures. For the sum scores derived from these instruments an ANOVA or ANCOVA was used for comparing means.

Dealing with missing data was somewhat challenging, as the reason behind the missing data is unknown. The data might be randomly missing (e.g. the parent forgot to answer the question), or it might be a specific reason for why the data is missing (e.g. the parent was uncertain how to answer or did not want to answer). Because of this uncertainty, data editing, replacing the missing value with the mean score (which is a traditional approach), might distort the data and make it biased (Schafer & Graham, 2002). In the current study missing data was dealt with by determining an acceptable number of missing items per scale, usually set to not more that 5% of the total number of items used to compose the scale, i.e. setting a "missing item threshold" for each variable. The threshold varies depending on the total number of items included in the variable. If the number of missing items exceeds the threshold, the case was excluded from the analysis. A more meaningful statistical method for dealing with missing data, such as maximum likelihood or multiple imputations, was considered but determined to be too methodological demanding for the scope of this thesis.

Outliers were investigated, however in most cases they were not excluded or recoded. It was expected that there would be a lot of variability in abilities across the sample, and excluding or recoding without knowing why they are outliners was considered problematic. However, for the VABS receptive scale three participants were excluded due to having extreme values (DQ > 200). Before excluding the participants they were compared to performance on other language measures (VABS expressive etc.) and found to have much lower DQ values. Therefor the high DQ score was considered a measurement error. This is probably a limitation with using age equivalents as foundations for a DQ scale (see section 3.4.2).

Lastly, it is worth noting that in tables were ANCOVA analysis was administered, the standard error (SE) will be presented instead of standard deviation (SD). The SE score is a measure of variability for the sample mean. It shows how much variability there is for the mean, across samples from the distribution. Larger SE values will indicate that the mean from the sample differs from the population (Field, 2009).

3.4 Validity and reliability

The following includes a review of the validity of the instruments, and the measurements constructed. It also includes an evaluation of potential participation bias. Some will characterize this as internal validity (Shadish, Cook, & Campbell, 2002).

3.4.1 Validity and reliability of the instruments

The term validity of an instrument refers to what extent the instrument measures what it purports to measure (Shadish et al., 2002). In the current study the validity of the instruments is well documented through their manuals, or other studies that have used the same measures. The instruments selected for use in the ABC-study are widely used in research of various populations, including both clinical and non-clinical samples. Most instruments are standardized and/or normed; however most of the instruments use American norms. This is a limitation because participants in Norway can interpret questions/tasks differently than American mothers/children. Some measures used in the ABC study have been translated specifically for use in the study. The recommended procedures for such instances have been followed (i.e. translation from English to Norwegian, "blind" back translation to English which then are reviewed by an independent professional). In addition, all discrepancies encountered are reviewed by two independent persons and resolved after discussions. For ADI-R the back-translated text has been reviewed by the publisher. Nevertheless, small differences in wording are impossible to avoid, and adaptations to Norwegian grammar and syntax is necessary.

Most of the language instruments used in the current study relies on parental reports. This has some limitations, but Ring and Fenson (2000) found good correlation between a parental checklist and direct assessment of the child's expressive abilities. Also Charman (2004) investigated ways of assessing pre-school children with ASD and he states that direct assessment might not be feasible, due to low language functioning. He also recommends combining the results from two different language measures, especially when using parental information. In the current study both CDI and VABS have an expressive language measure. Both rely on parental reports, but one is a checklist for the parent while the other is interview-based.

To ensure reliable use of the instruments the staff went through extensive training prior to using the instruments. As quality assurance each research assistant had to document at least 90 % agreement on three consecutive ADI-R and VABS interviews that were conducted by a senior research assistant, prior to being able to conduct interviews alone. The psychologists were mentored by a senior clinician in learning the instruments and there were weekly case reviews for verification of diagnoses. Monthly staff-meetings were held to assure ongoing reliable use of the instruments, especially ADI-R and ADOS.

3.4.2 Validity and reliability of the constructed measures

The cognitive composite measures (FISQ and NVIQ) used in this study have limitations that might influence their validity. They are constructed from items selected on the basis of different theoretical perspective (Mullen and SB5). The standardized IQ from SB5 reflects cognitive functioning with assessments that measure working memory, abstract reasoning etc. However, the Mullen measure reflect general development by using sub-test of fine motor skills, visual reception and expressive and receptive language. Hence they are measuring two different phenomena, which is a possible threat to the construct validity of the composite measure (Shadish et al., 2002). However, the Mullen measure was only used when the child was unable to complete the SB5, and it was therefore considered a necessary trade-off to use the Mullen as an indicator of the child's cognitive functioning. Another potential threat to the validity of this measure is the non-standardized conversion of the "Mullen T-score" to get a full scale IQ, as well as using both the "SB5 estimated NV" and the "Mullen converted NV" variables. While their construction has good face validity, they lack normative data from direct testing across age groups. It is worth noting that no measures were constructed by using sub-tests from Mullen and SB5 in one combined measure, both were constructed only using items from each of the individual test.

A note should also be made on the SB5 abbreviated IQ measure. Coolican, Bryson, and Zwaigenbaum (2008) found the abbreviated IQ to be representative of the full-scale measure. However they did find the abbreviated IQ measure had a tendency to overestimate the child's abilities compared to the full-scale measure. They attributed this difference to the different weighing of the subtests (i.e. two subtests in the abbreviated compared to ten in the full-scale), especially the Fluid Reasoning subtest, which accounted for much of the variation.

Another potential limitation is the use of age equivalents for constructing the VABS DQ and CDI DQ measures. Mervis and Klein-Tasman (2004) claim that the age equivalent score is highly problematic because it is derived from the median chronological age at which a raw score is obtained, and thus it is not an interval scale. The 4 month difference between 16 and 20 months, and 36 and 40 months should represent the same amount of growth. However this is not the case.

Lastly there has also been discussion in the literature regarding using IQ as a covariate in analysis. In the current study NVIQ was used as a covariate for all ANCOVA analysis. Dennis et al. (2009) argues that IQ is not an appropriate covariate in studies of neurodevelopmental disorders (like ASD and LI). They argue that IQ should be considered a measure of achievement rather than an indication of potential. They also say it does not meet the requirements of a covariate, because the relationship it has to the outcome and independent variable is complicated (e.g. the neurodevelopmental disorder precedes the IQ score, and can therefore never be separated from the effects of the disorder). Therefore using IQ as a covariate will produce overcorrected findings in ANCOVA analysis, and making the groups seem more similar than they actually are. Regardless, NVIQ is often used as cofounder in research on group differences and therefore considered appropriate for the current study. However the objections raised by Dennis et al. (2009) regarding the use of NVIQ should be noted.

3.4.3 Potential participation-bias

Participation in the ABC-study, and consequently the current study, is related to who participates in the MoBa study. Nilsen et al. (2009) investigated potential participation-bias in the MoBa cohort, and found the study had an underrepresentation of women under 25, of those living alone, mothers with more than two previous births and with previous stillbirths. In addition it was found that smokers were underrepresented, and women taking health supplements were overrepresented.

In addition to the selection bias in MoBa, there might be a bias regarding who of those invited to the clinical assessment could partake in the ABC-study. An article investigating possible participation-bias in the ABC-study is submitted for publication (Nielsen, submitted). One potential participation-bias might have been those living far from the assessment-site that would be unable to participate in an assessment in Oslo, due to having a child with high levels

of difficulties making travelling a challenge. The ABC-screening criteria might also result in missing participants if the criteria produce many false negatives (i.e. the instrument wrongly indicating not having ASD). If this is the case the assessed sample will not be representative for the total variability in ASD symptomatology. However, the criterion was built on the current best-estimate of how to screen for autism in young children.

3.5 Ethical considerations

Both the ABC-study and MoBa are approved by the Regional Committees for Medical and Health Research Ethics (REK) in Norway. Participation in the ABC-study requires an additional informed consent to the parent that consented to participation in MoBa. Participation is voluntary, and the participants can at any time, without giving a reason, leave the study. If they leave they can have all data and samples related to them erased. All data collected in the ABC study are de-identified before any researchers are given access to the data. Key variables are omitted or recoded so as to not leading to a possible identification of individual families/ children. As an example, the birth date or assessment date is never revealed in the dataset, appearing only as recalculated age in days when being assessed in the ABC study. Data used in the current thesis has been de-identified. In accordance with NIPH policy all statistical analyses have been done with data stored within the institutions computer network.

Many families who came to an assessment with severely affected children were in need of advice and time to discuss their challenges with professionals. The ABC clinic had skilled professionals, who also wrote reports stating the needs of the child, and how to get in touch with the appropriate local services. This was an important aspect of running the ABC-study, as many participating children and families had unmet needs. Seeing so many children with ASD, LI and various other difficulties, made it possible to help local services reach a diagnostic conclusion and starting appropriate interventions during pre-school.

Regardless of being approved by the ethics committee, research on children and especially marginalized groups such as children with ASD and LI requires special attention to ethical considerations. On a more philosophical level, one of the main discussions in the literature is to what extent there are differences between a child and an adult participating in research. According to Morrow and Richards (1996) this debate has been characterized through the

adults perception of what a child needs. The main worry is that the child can be indirectly misled or pressured by the researcher, and therefore needs protection and safekeeping. This view has been debated, as societies need for knowledge about children's development has increased and their experiences are considered important. This might, however, come in conflict with the child's right to protection. It also been proposed that children have the right to be heard, and consequently also have the right to participate in research (Backe-Hansen, 2012). The current practice with parents consenting on behalf of their child has been problematized. Coyne (2010) states that parental consent can result in the child's rights being ignored or not receiving the attention it should get. On the other hand, the use of child consent in studies with pre-school children does not seem to be commonly used in research today, and how this can be accomplished in a way the child understands is also debatable.

Relating this to the current study, the ABC-study protected the children from unnecessary stress during the assessment by following each child's needs and signals. The staff also had extensive experience assessing children making them sensitive to the children's needs. The children's privacy was also protected by ensuring that data was de-identified. The ABC-study uses parental consents only.

4 Results

In the following section the results from the statistical analyses will be presented. The mean and standard error/standard deviation results from the separate analysis run on gender are presented in appendix B (boys) and C (girls).

4.1 Distribution of variables

Table 3. *Distribution*; *Sample characteristic by variable*

	Group N							
Variable	ASD (N=54)	LI (N=161)	TD (N=414)	Total N	Total mean	Total SD	Total Skew	Total Krt
IQ composite scores								
FSIQ	54	161	414	629	98.7	15.1	-.399	-.262
NVIQ	54	160	414	628	102.6	14.0	-.522	.210
CDI factors								
CDI DQ	49	138	318	505	87.8	27.1	.906	.770
Length of utterance	50	141	330	521	2.1	1.1	-.843	-.729
Vocabulary	50	138	330	518	3.6	0.8	-2.258	5.495
Grammar	50	141	329	520	1.5	1.1	-.080	-1.331
Pronouns	49	138	329	516	2.9	2.9	-0.954	-0.163
VABS domains								
Expressive DQ	52	158	401	611	90.4	28.0	.558	1.821
Receptive DQ	52	158	399	609	94.4	22.6	-.401	-.453
PPBS factors								
Social play	20	54	193	267	23.5	5.0	-.851	.325
Solitary - passive	21	65	206	292	15.4	2.6	-.044	-.542
ADI-R algorithm domains								
Communication (V)	40	118	277	435	3.1	3.9	1.890	3.944
Social (V)	40	118	277	435	2.2	3.5	2.817	10.389
Repetitive (V)	40	118	275	433	1.1	2.1	2.933	10.227
Communication (NV)	14	39	-	53	4.0	3.3	.588	-.594
Social (NV)	14	39	-	53	5.6	1.3	1.302	1.281
Repetitive (NV)	14	39	-	53	1.8	0.7	.709	-.817

Note. Krt = Kurtosis V = verbal participants, NV = non-verbal participants, as defined by the ADI-R.

Number of participants

There are 54 participants in the ASD group, 161 in the LI group and 414 in the TD group, though there is some variation in the number of participants across instruments. This is due to changes in the study protocol during the course of the ABC-study. Firstly, the number of typically developing children who have been assessed with the ADI-R is lower than the number that has been seen for a clinical assessment. This is related to a change in protocol early in the study where children selected as controls were not administered the ADI-R. Secondly, the PPBS was only included in the pre-assessment form for the parents up until 2009; from then onward it was only included in the pre-assessment forms sent to the daycare, to be filled in by the pre-school teacher. The rest of the instruments have higher numbers of participants since these were administered to all participants throughout the study. The other variation in numbers is related to exceeding the missing data thresholds (i.e. number of missing items exceeds an acceptable limit for that variable, usually 5%, or completely missing data when the instrument was never administered to the subject for different reasons.

Gender distribution

There are substantial gender differences (male:female) in both the ASD group (78.3 % males and 21.7 % females) and the LI group (78.3 % males and 21.7 % females), while there is an equal gender distribution in the TD group (52.4 % males and 47.6 % females).

Table 4 shows the number of participants for the main groups of variables, split on gender. It is worth noting that some of the variables have very few girls, especially in the ASD group.

Table 4.
Gender; Number of participants by gender for main variables

Variable	ASD boy/girl	LI boy/girl	TD boy/girl
IQ composite scores	47/7	125/35	217/197
CDI factors	45/5	109/32	169/160
VABS domains	45/7	123/35	210/190
PPBS factors	16/4	43/11	97/98
ADI-R algorithm domains (V)	35/5	90/28	152/125
ADI-R algorithm domains (NV)	12/2	33/6	-

Skewness and kurtosis

To investigate the distribution of variables, values for skewness and kurtosis were calculated. Skewness refers to the symmetry and kurtosis to the pointiness of the distribution. Both values should be zero if the distribution is perfectly normally distributed. Both measures can have a positive and negative prefix, thus mirroring the direction of how the distribution differs from normality. The further the values are from zero the more likely it is that the distribution is not normal. It seems problematic to establish a critical value for when the values of skewness and kurtosis are suboptimal. It has been noted that large sample sizes can affect the values in a negative way without posing a threat to the validity of the analysis (Field, 2009). Also, Pallant (2010) states that in samples with over 30 participants, violation of this assumption should not cause major problems. However Cameron (2004) proposes that values should fall in the range from +2 to -2 for a normal distribution. For the current sample most values fall within this range, except for the ADI-R algorithm variables which have high skew and kurtosis values. This is, however, as expected as the TD and LI group should have low scores on this instrument (i.e. they would cluster at the low end of the scale). Additionally the CDI Vocabulary had high skew and kurtosis variables, reflecting a ceiling effect on the variable.

Since the current study is comparing different groups, matching skew and kurtosis across the whole sample has limited value for many of the variables, since the three groups in the current sample are very different. It is more important that each group has a normal distribution for each variable (Field, 2009). When comparing the scores by each individual group most fall within the range proposed by Cameron (2004), except for the same ADI-R variables as mentioned for the whole sample. However the ADI-R verbal communication sub-domains seem fairly normally distributed for the LI group in addition to the ASD group.

4.2 Sample characteristic

Table 5 shows number of participants within each FSIQ group cross tabulated with diagnostic group. The table shows that there are 29 participants in the ASD group that has an IQ lower than 54. There are no participants in the LI and TD groups with such low scores, thus it is difficult to compare them with children in the other diagnostic groups. These 29 participants are therefore excluded from further comparative analysis. However, they will be described more fully in a separate section.

Table 5.
FSIQ; Cognitive functioning distribution across groups

IQ group	ASD group	LI group	TD group
< 54	29	0	0
55-69	14	8	2
70-84	16	68	17
85-99	10	56	105
100-114	11	27	206
115 <	3	2	84

It is also worth noting that there are two participants from the TD group scoring within the 55 – 69 FSIQ span, leading to suspicion of inconsistencies in the data quality or erroneous diagnostic conclusions. However examining the NVIQ scores they all score within the 70 – 84 range which showed a large discrepancy compared to the VIQ. It was considered that the conclusion that these children were not intellectual impaired is valid. They were not excluded from the analysis.

Table 6.
Sample demographics; age, FSIQ and NVIQ

	ASD group *(N=54)*			LI group *(N=161)*			TD group *(N=414)*			ANOVA	
	mean	SE	95% CI	mean	SE	95% CI	mean	SE	95% CI	*F*	P-H[a]
Age (months)	42.6	0.6	41.4-43.8	41.1	0.2	40.8-41.5	42.3	0.1	42.1-42.6	-	-
FSIQ[a]	83.0	2.4	78.8-88.4	87.1	0.9	85.3-88.9	105.2	0.6	104.0-106.3	174.7*	A=L<T[b]
NVIQ[a]	86.8	2.3	82.1-91.5	94.3	0.9	92.5-96.1	107.8	0.5	106.7-108.9	128.5*	A<L<T[c]

*$p < .001$
Notes:
[a] P-H: Games-Howell post hoc test (A=ASD, L=LI, T=TD)
[b]The ASD group was not significantly different from the LI group. But both are different from the TD group
[c]The ASD group were significantly lower than the LI group and TD group. The LI group was significantly lower than the TD group, but significanltly higher than the ASD group

Sample characheristics of age, FSIQ and NVIQ across diagnostic groups is shown in table 6.

The ANOVA for FSIQ shows significant differences between the groups $F(2, 626) = 174.7$, $p < .001$, $\eta^2 = .358$. The Game-Howell post hoc test shows there is no significant difference between the LI and ASD group for FSIQ. On the NVIQ variable, the ANOVA shows a significant difference $F(2, 625)$, $p < .001$, $\eta^2 = .358$. The Game-Howell post hoc test shows significant differences between all three groups.

When running the ANOVA post-hoc analysis by gender there is no significant differences between LI and ASD on FSIQ, and significant differences between all groups for NVIQ for boys. However for girls there were no significant differences between ASD and LI either for NVIQ or FSIQ. For sample demographics separately on gender see appendix A.

4.3 Language milestones

Table 7.
Language milestones; age of first word and age of first phrase

	ASD group			LI group			TD group			ANOVA	
	N	mean	SD	N	mean	SD	N	mean	SD	*F*	P-H[a]
Age first word (months)	51	17.2	7.5	142	19.5	4.3	244	14.7	4.3	31.1[*]	A=L<T[b]
Normal, months unknown	1			7			27				
Delayed, months unknown	0			3			1				
Milestone not achieved	0			1			0				
Age of first phrase (months)	52	25.7	8.0	146	29.7	6.3	275	20.4	4.6	132.6[*]	L<A<T[c]
Normal, months unknown	1			7			41				
Delayed, months unknown	0			0			1				
Milestone not achieved	2			8			0				

[*] $p < .001$
Notes:
[a]P-H: = Games-Howell post hoc test (A=ASD, L=LI, T=TD).
[b]The ASD group were not significantly different from the LI group. Both the LI and ASD group was significantly lower than the TD group
[c]The LI group were significantly lower than the ASD group and TD group. The ASD group was significantly lower than the TD group, but significantly higher than the LI group

Table 7 shows language milestones across diagnostic groups. The mean age of first word and first phrase was calculated from the ADI-R. Some of the participants did not have a month value, but had a value indicating a best-estimate code (1- normal, months unknown, 2 - delayed, months unknown, 3 - milestone not achieved). These participants were counted, but naturally a mean and SD score was not calculated. The ADI-R defines adequate acquisition of first word prior to 24 months, and prior to 33 months for age of first phrase.

The mean age for first word for the ASD group was 17.2 months. In addition there were 11 participants who had not achieved this milestone. The LI group had a mean age of 19.5 months, with one participant not having archived the milestone. The TD group had a mean age of 14.7.

Running an ANOVA to compare differences on first word acquisition shows a significant difference between means, $F(2, 436) = 31.107$, $p < .001$, $\eta^2 = .125$. The post hoc analysis

reveals no significant difference between the ASD and LI groups. However, they were both different from the TD group. The relatively high SD reveals great variation in acquisition of first word.

Running the ANOVA for first phrase also shows a significant difference between means, $F(2, 472) = 132.6$, $p < .001$, $\eta^2 = 360$. For this variable the post hoc test shows significant differences between all groups. The LI group had the highest mean value, showing perhaps the greatest delay. Also both the LI and the ASD group had a high SD score, showing great variability in acquisition.

Running the ANOVA post hoc separately for both genders showed there was a significant difference between LI relative to both ASD and TD on acquisition of first word for boys. There was no significant difference between ASD and TD. This means the LI group was the most delayed. For age of first phrase there were significant differences between all groups. For girls there was no significant difference between ASD relative to both LI and TD for age of first word. However there was a significant difference between TD and LI. The same patter existed for age of first phrase. For mean values from the analysis see appendix B and C.

4.4 Language functions

In the following section both results from the complete CDI DQ measure and the various language factors will be presented.

4.4.1 CDI expressive DQ scale

The CDI expressive language score was standardized as a DQ. The scale met the assumption of homogeneity of regressions slopes, an ANCOVA was therefore considered suitable. The measure shows mean values of ASD = 70.7, LI = 67.4 and TD = 99.3. There was significant differences between the groups, $F(3, 501) = 144.86$, $p < .001$, $\eta^2 = .465$. The covariate NVIQ was significantly related to the CDI score, $F(1, 501) = 24.7$, $p < .001$, $\eta^2 = .047$. NVIQ explained about 4.7 % of the total variance. There was also a significant diagnostic group effect after controlling for NVIQ, $F(2, 501) = 102.4$, $p < .001$, $\eta^2 = .290$. Diagnostic group explained 29.0 % of the total variance of the language score.

It is worth noting that the CDI DQ score was very low for the LI group, in spite of an adequate NVIQ. The post-hoc test shows that there is no significant difference between ASD and LI on the CDI-expressive language total score, but they were both significantly different from the TD group. The VABS expressive DQ and CDI DQ were significantly correlated, r = .735, p (two-tailed) < .001.

When running separate analysis on gender, the differences between the diagnostic groups persisted. For mean values from the analysis on gender see appendix B and C.

4.4.2 CDI language factors

Various factors were extracted (see section 3.3.2) from the CDI, and group differences at 3 and 4 years were examined. Note that children at age 3 and at age 4 are independent samples (i.e. not the same participants seen two times), and the percentages should act accumulative (e.g. a participant that scores >10 word, will probably also score positive on > 6 words and > 4 words). Each item is split into participants that master the language function at three years (between 3 years and 0 months - 3 years and 11 months), and at four years (between 4 years and 0 months - 4 years and 12 months). In addition, each factor was coded into a sum score on an interval level, and analyzed with an ANCOVA. The wording of the questions has been

shortened to fit inside the table; for a complete overview of the full wording of each question see section 3.3.2.

Length of utterances

Table 8.
CDI; Length of utterance, number and percentages of participants mastering the language function

	ASD group		LI group		TD group	
	n	%	n	%	n	%
3 year old sample[a]						
Talks in sentences > 4 words	17	63.0	55	56.7	170	99.4
Longer > 6 words	9	33.3	23	23.5	166	96.5
Long, complex > 10 words	4	14.8	2	2.1	121	72.5
4 year old sample[b]						
Talks in sentences > 4 words	14	60.9	32	74.4	158	100.0
Longer > 6 words	9	39.1	17	39.5	153	96.2
Long, complex > 10 words	4	18.2	6	14.0	135	88.2

[a]Participants 3.0 - 3.12 : ASD N=27, LI N=98, TD N=172

[b]Participants 4.0 - 4.12: ASD N=23, LI N=43, TD N=159

Table 8 shows the results for the length of utterance. The category within each age group reflects using longer and more complex utterances. Looking at three year olds across diagnostic groups, 37 % of the children with ASD use shorter sentences than four words, 43.3 % of the children with LI use less than four words and only 0.6% of TD use less than 4 words in their sentences. The results show that as length and complexity increases on this measure the rate of participants mastering the function in the ASD and LI group drops sharply, while the percentages of participants mastering the skill in the TD group remains high. The three measures of length of sentences were summarized into a composite scale ranging from 0=not mastering any item to 3= mastering all items. The composite measure show mean values for ASD= 1.1; LI= .9 and TD= 2.7. Since the assumption of homogeneity of regression slopes were not fulfilled an ANCOVA was not considered suitable.

Since an ANCOVA was not suitable, an ANOVA was used to look at differences across groups. The analysis showed a significant difference between groups, $F(2, 518) = 359.9$, $p < .001$, $\eta^2 = .580$. The Game-Howell post hoc test shows there is no significant difference between the LI and ASD groups, but they were significantly different from the TD group.

When running the post-hoc analysis separately for each gender, there was no significant difference between the ASD and LI groups, for either boys or girls. However both remained significantly different from the TD group. For mean values from the analysis see appendix B and C.

Vocabulary size

Table 9.
CDI; Expressive vocabulary size, number and percentages of participants mastering the language function

	ASD group		LI group		TD group	
	n	%	n	%	n	%
3 year old sample[a]						
> 2 words other than "mama"	26	96.3	91	94.8	172	100.0
Uses at least 10 words	23	85.2	82	87.2	172	100.0
Vocabulary ≥ 20	25	92.6	84	88.4	170	99.4
Uses ≥ 50 different words	17	65.4	34	35.4	156	94.0
4 year old sample[b]						
> 2 words other than "mama"	21	91.3	40	93.0	159	100.0
Uses at least 10 words	21	91.3	36	83.7	157	99.4
Vocabulary ≥ 20	20	87.0	35	81.4	156	98.1
Uses ≥ 50 different words	13	59.1	24	57.1	145	94.2

[a]Participants 3.0 - 3.12 : ASD N=27, LI N=98, TD N=172

[b]Participants 4.0 - 4.12: ASD N=23, LI N=43, TD N=159

Table 9 shows vocabulary size. The categories listed reflect increased size of vocabulary running from "> 2 words other than "mama" and up to "uses > 50 different words". The table shows results grouped by age. The results shows that a very high percentage in each

diagnostic group master Vocabulary $\geq$ 20 words. For the highest category (uses > 50 different words) both the ASD and LI group show a sharp drop in percentages of children who master this for both 3 and 4 year olds. For 3 year olds the LI group scores lower than the ASD group.

The four items were then summarized into a composite scale ranging from 0= not mastering any items to 4= mastering all items. The scores in the diagnostic groups showed mean values for ASD= 3.3, LI= 3.0 and TD= 3.8. The assumption of homogeneity of regression slopes were fulfilled, therefore an ANCOVA was considered suitable. There was significant difference between the groups, $F(3, 513) = 59.657$, $p < .001$, $\eta^2 = .259$. The covariate was significantly related to the vocabulary score, $F(1, 513) = 7.439$, $p = .007$, $\eta^2 = .014$. NVIQ explained 1.4 % of the total variance. Diagnostic grouping was also significantly related to vocabulary, $F(2, 513)$, $p < .001$, $\eta^2 = .162$. It explained 16.2 % of the total variance. The Bonferroni post hoc test revealed significant differences between all groups.

When running the analysis separately on gender there was significant differences for all groups for boys, but no significant difference between ASD and LI for girls. For mean values see appendix B and C.

Grammar skills

For grammar skills, the items have been sorted in descending order based on the percentage of mastery in the TD group.

Table 10.
CDI; Grammar skills, number and percentages of participants mastering the language function

	ASD group		LI group		TD group	
	n	%	n	%	n	%
3 year old sample[a]						
Uses plural words, adding "s"	10	38.5	21	21.9	153	90.0
Talks in past-tense correctly	8	29.6	9	9.2	124	73.4
Uses plurals correctly, e.g."men" not "mans"	6	22.2	4	4.1	54	32.9
4 year old sample[b]						
Uses plural words, adding "s"	10	43.5	22	52.4	145	94.2
Talks in past-tense correctly	8	36.4	13	31.7	130	82.3
Uses plurals correctly, e.g."men" not "mans"	2	8.7	5	11.6	63	41.4

[a]Participants 3.0 - 3.12 : ASD N=27, LI N=98, TD N=172

[b]Participants 4.0 - 4.12: ASD N=23, LI N=43, TD N=159

Table 10 shows grammar skills. As expected, the percentages of children in the ASD and LI group mastering each of the grammar skills are much lower than typically developing children, reflecting difficulties with mastering grammar both in 3 and 4 year olds. The 4 year olds score higher than the three year olds on all items, showing a developmental trend.

The three items were then summarized into a composite scale ranging from 0 = not mastering any items to 4 = mastering all items. The scores in the diagnostic groups show mean values for ASD= 1.1, LI= .6 and TD= 1.9). The assumption of homogeneity of regression slopes was fulfilled, so an ANCOVA was suitable. There was significant difference between groups, $F(3, 515) = 118.0$, $p < .001$, $\eta^2 = .407$. The covariate NVIQ was significantly related to the grammar score, $F(1, 515) = 19.453$, $p < .001$, $\eta^2 = .036$. NVIQ explained 3.6 % of the variance. There was also a significant diagnostic group effect after controlling for NVIQ, $F(2, 515) = 89.678$, $p < .001$, $\eta^2 = .258$. Diagnostic group explained 25.8 % of the total

variance. The Bonferroni post-hoc test shows there is a significant difference between all groups, with the LI group being the most impaired.

When running the post-hoc test separately for gender there was significant differences between all groups for boys, however, there were no significant differences between ASD and LI for girls. They were significantly different from the TD group. For mean values see appendix B and C.

Pronouns

Table 11 shows pronouns. The items measuring use of pronouns have been ordered according to descending scores by the TD group. The percentages of children mastering each item show a similar trend in the ASD and LI group, as in the TD group, indicating a similar developmental trend across all groups. A higher percentage of the LI group scored better than the ASD group on all items for both age groups.

Table 11.
CDI; Pronouns, numbers and percentages of participants mastering the language function

	ASD group		LI group		TD group	
	n	%	n	%	n	%
3 year old sample [a]						
Refers to thing as "my" or "mine"	16	59.3	79	80.6	170	99.4
Uses the word "you" in sentences	11	40.7	48	49.0	165	98.2
Uses the words "me", "my" and "I" correctly	8	34.8	41	47.7	140	88.1
Uses plural pronouns such as "we", "you"	4	15.4	18	18.8	118	70.2
4 year old sample [b]						
Refers to thing as "my" or "mine"	17	73.9	37	90.2	157	98.7
Uses the word "you" in sentences	13	56.5	26	61.9	158	99.4
Uses the words "me", "my" and "I" correctly	8	38.1	26	65.0	129	94.9
Uses plural pronouns such as "we", "you"	7	31.8	8	19.5	121	79.1

[a]Participants 3.0 - 3.12 : ASD N=27, LI N=98, TD N=172

[b]Participants 4.0 - 4.12: ASD N=23, LI N=43, TD N=159

The four measures of pronouns were summarized into a composite scale ranging from 0=not mastering any item to 4= mastering all items. The scale met the assumption of homogeneity of regression slopes, so an ANCOVA was considered suitable. The analysis generated mean values for all groups (ASD: 1.7, LI: 2.0, TD: 3.5). There was significant difference between groups, $F(3, 512) = 123.6$, $p < .001$, $\eta^2 = .420$. The covariate NVIQ was significantly related to the pronoun score, $F(1, 512) = 23.6$, $p < .001$, $\eta^2 = .044$. NVIQ explained 4.4 % of the variance. There was also a significant diagnostic group effect after controlling for NVIQ, $F(2, 512) = 71.7$, $p < .001$, $\eta^2 = .243$. Diagnostic group explained 24.3 % of the total variance. The Bonferroni post-hoc test shows there is no significant difference between the ASD and LI group, while they were significantly different from the TD group.

When running the post-hoc test separately on gender there was no significant differences between ASD and LI for boys, but there was a significant difference between ASD and LI for girls. However they were both groups were different from the TD group, regardless of gender. For mean values see appendix B and C.

Narratives

Table 12.

CDI; Narratives, numbers and percentages of participants mastering the language function

	ASD group		LI group		TD group	
	n	%	n	%	n	%
3 year old sample [a]						
Tells you what's going on in a picture	6	27.3	5	12.2	115	74.2
Retells short stories, tells what happens in correct order	2	7.4	3	3.1	91	53.5
4 year old sample [b]						
Tells you what's going on in a picture	19	82.6	37	86.0	158	100.0
Retells short stories, tells what happens in correct order	20	74.1	78	81.2	169	98.8

[a]Participants 3.0 - 3.12 : ASD N=27, LI N=98, TD N=172

[b]Participants 4.0 - 4.12: ASD N=23, LI N=43, TD N=159

Table 12 shows narratives. The items measuring use of narratives have been ordered according to descending scores by the TD group.

All groups had low scores on these variables at three years. However, in the four year sample, almost the entire TD group mastered the items, and the LI and ASD group had increased significantly compared with the three year olds. Since the scale has a low Cronbach's alpha value, and has few items, an ANOVA was considered unfitting.

4.5 Language qualities

In this section atypical language from the ADI-R and language factors from CDI will be presented.

4.5.1 ADI-R: Atypical language

Atypical language qualities were measured by comparing a number of language items from the ADI-R. Table 13 shows the percentage of participants in each group, across different scores on each item. The general pattern across the items is that the ASD group has the highest scores, reflecting higher levels of impairment. While the LI group scores higher than the TD group, they have lower scores than the ASD group. This reflects either lower impairment in the LI group relative to the ASD group, or uncertainty in the quality of the behavior.

Table 13.

ADI-R: atypical language questions

		ASD group *(N=54)*		LI group *(N=157)*		TD group *(N=277)*	
		n	%	n	%	n	%
Overall level of language	0 = functional use	40	74.1	118	75.2	277	100.0
	1 = no functional use	12	22.2	36	22.9	-	-
	2 = fewer than five words	2	3.7	3	1.9	-	-
Reciprocal conversation	0 = conversation flows	8	19.0	45	38.1	255	92.4
	1 = occasional conversation	16	38.1	37	31.4	20	7.2
	2 = little or no conversation	17	40.5	36	30.5	1	0.4
	3 = very little speech	1	2.4	-	-	-	-
Inappropriate questions or statements	0 = no or very rare	36	87.8	111	94.9	240	87.0
	1 = sometimes	5	12.2	5	4.3	35	12.7
	2 = frequent	-	-	1	0.9	1	0.4
Pronoun reversals	0 = no confusion	15	36.6	70	64.2	245	89.7
	1 = uses name, but not other	16	39.0	30	27.5	23	8.4
	2 = you/I or s/he/I, without question intonation	5	12.2	7	6.4	5	1.8
	3 = same as 2, but with question intonation	5	12.2	2	1.8	-	-
Verbal rituals	0 = nothing	30	69.8	111	92.5	271	97.8
	1 = tendencies	8	18.6	6	5.0	5	1.8
	2 = one or more rituals	3	7.0	3	2.5	1	0.4
	3 = as 2, but greater social impairment	2	4.7	-	-	-	-
Intonation/volume/rhythm /rate	0 = normal	26	61.9	99	84.6	253	91.3
	1 = some, but doesn't affect intelligibility	12	28.6	12	10.3	19	6.9
	2 = clearly abnormal, affects intelligibility somewhat	4	9.5	5	4.3	5	1.8
	3 = obvious abnormal	-	-	1	0.9	-	-
Current communicative speech	0 = used frequently	16	38.1	80	66.1	269	97.1
	1 = restricted communication	22	52.4	35	28.9	8	2.9
	2 = spontaneous words	4	9.5	5	4.1	-	-
	3 = little or no communication	-	-	1	0.8	-	-

4.5.2 VABS: Language used in everyday living

The first language comparison is done using the Vineland receptive and expressive subdomain DQs. Analysis for the three measures was done using ANCOVA, controlling for nonverbal IQ.

Table 14.
VABS; Language used in everyday living, adjusted for NVIQ

	ASD group *(N=52)*			LI group *(N=158)*			TD group *(N=401)*			ANCOVA	
	mean	SE	95% CI	mean	SE	95% CI	mean	SE	95% CI	*F*	P-H[a]
Vineland DQ											
Receptive	76.4	2.9	70.7-82.1	88.5	1.6	85.3-91.7	99.0	1.0	97.0-101.0	75.9[*]	A<L<T[b]
Expressive	72.5	3.1	66.4-78.7	67.9	1.7	64.5-71.4	101.6	1.1	99.5-103.8	168.4[*]	A=L<T[c]

[*] $p < .001$

Notes:

[a]P-H: Bonferroni post hoc test (A=ASD, L=LI, T=TD)

[b]The ASD group were significantly lower than the LI group and TD group. The LI group was significantly lower than the TD group, but significantly higher than the ASD group.

[c]The ASD and LI groups were significant lower than the TD group.

Table 14 shows language used in everyday living, as measured by Vineland.

The Vineland receptive DQ shows a significant difference between ASD, LI and TD groups. The covariate NVIQ was significantly related to the VABS score, F (1, 605) = 46.2, $p < .001$, $\eta^2 = .071$. NVIQ explained 7.1 % of the total variance. There was also a significant group effect after controlling for NVIQ, F (2, 605) = 29.1, $p < .001$, $\eta^2 = .088$. Diagnostic group explained 8.8 % of the total variance. When running separate analyses for gender there were no difference between the LI group and TD group for girls. However, there was still a significant difference for boys. For mean values from the analysis see appendix B and C.

The Vineland expressive DQ and the covariate NVIQ was significantly related to the VABS score, F (1, 607) = 29.2, $p < .001$, $\eta^2 = .046$. NVIQ explained 4.6 % of the variance. There was also a significant diagnostic group effect after controlling for NVIQ, F (2, 607) = 126.0, $p < .001$, $\eta^2 = .293$. Diagnostic group explained 29.3 % of the total variance. When running separate analyses for gender, the same differences between the groups persisted. For mean values see appendix B and C.

In plain words, there was no significant difference between ASD and LI in terms of expressive language, while the ASD group was more impaired than the LI group on receptive language.

4.6 PPBS: Play behavior

Table 15 shows the mean scores on both the solitary – passive and social play factor. The solitary – passive factor shows that the covariate, NVIQ, was not significantly related, $F(1, 287) = .242$, $p < 0.623$, $\eta^2 = .001$. Likewise, there was no significant effect of the grouping variable, $F(2, 287) = 2.1$, $p < .112$, $\eta^2 = 0.015$. The post hoc test shows no differences between the groups, and running the analysis separately for gender shows the same. For the current sample this factor shows excellent internal consistency of .917, as measured by Cronbach's Alpha.

Table 15.
PPBS; play behavior, adjusted for NVIQ

	ASD group (N=21)			LI group (N=53)			TD group (N=206)			ANCOVA	
	mean	SE	95% CI	mean	SE	95% CI	mean	SE	95% CI	*F*	P-H[a]
Solitary - passive	16.4	0.6	15.2-17.6	15.1	0.2	14.4-15.7	15.4	0.2	15.0-15.7	1.4*	A=L=T[b]
Social play	16.6	0.9	14.8-18.5	21.0	0.5	20.0-22.2	24.8	0.3	24.3-25.4	66.6**	A<L<T[c]

* p .234

** p < .001

Notes:

[a]P-H: = Bonferroni post hoc test (A=ASD, L=LI, T=TD)

[b]No significant differences between the ASD, LI and TD groups

[c]Signicant differences between the ASD, LI and TD group. The ASD group were significantly lower than the LI group and TD group. The LI group was significantly lower than the TD group, but significantly higher than the ASD group

The social play factor was significantly related to the covariate nonverbal IQ, $F(1, 262) = 16.0$, $p < .001$, $\eta^2 = .058$. Nonverbal IQ explains 5.8 % of the total variation. There was also a significant effect of the group variable, $F(2 , 262) = 38.8$, $p < .001$, $\eta^2 = 0.228$. This variable explains 22.8 % of the total variation. The post hoc test shows significant differences between all groups.

Running the analysis separately for gender, it shows the same pattern for boys, however for girls there was no significant difference between ASD and LI. For mean values see appendix B and C.

4.7 Autism Symptomatology

The autism symptomatology was measured using selected items from the ADI-R. The communication domain in the interview is scored differently depending on whether the child's language abilities were considered verbal or nonverbal. The ADI-R defines a subject as verbal if the child uses utterances of at least three words on a daily basis. For the ASD group 40 participants were in the verbal group, while 14 were scored as nonverbal. For the LI group 118 were scored as verbal, while 39 were scored as nonverbal. All of the TD participants were scored as verbal.

Table 16.
ADI-R; algorithm scores for verbal participants, adjusted for NVIQ

	ASD group *(N=40)*			LI group *(N=118)*			TD group *(N=277)*			ANCOVA	
	mean	SE	95% CI	mean	SE	95% CI	mean	SE	95% CI	*F*	P-H[a]
Communication (V)	9.4	0.5	8.4-10.3	4.5	0.3	4.0-5.0	1.6	0.2	1.2-1.9	131.8[*]	A<L<T[b]
Social	9.1	0.4	8.2-10.0	2.2	0.3	1.7-2.7	1.2	0.2	0.9-1.6	111.6[*]	A<L<T[b]
Repetitive	4.8	0.3	4.2-5.3	1.1	0.2	0.8-1.4	0.6	0.1	0.4-0.8	83.0[*]	A<L<T[b]

[*] $p < .001$
Notes:
[a]P-H: = Bonferroni post hoc test (A=ASD, L=LI, T=TD)
[b]The ASD group was significantly lower than the LI group and TD group. The LI group was significantly lower than the TD group, but significantly higher than the ASD group

Table 16 shows the ADI-R algorithm scores for verbal participants. An ANCOVA was run to determine the group differences on all ADI-R algorithm variables, controlling for NVIQ. For verbal children, the covariate NVIQ was significantly associated with communication, $F(1, 431)$, = 11.8, $p < .001$, $\eta^2 = .027$ and accounted for 2.7 % of the total variance. Diagnostic group was significantly related to the presence of abnormality in communication, $F(2, 431)$ =

120.1, $p < .001$, $\eta^2 = .358$. Diagnostic group explained 35.8 % of the total variance, controlling for NVIQ. The post hoc test shows significant differences between all groups.

For the social domain NVIQ was not significantly related, $F(1, 431) = 2.28$, $p < .131$, $\eta^2 = .005$. However, diagnostic group was significantly related, $F(2, 431)$, $p < .001$, $\eta^2 = .374$. This explained 37.4 % of the total variance. The post hoc test shows that there are significant differences between all groups.

Lastly, on the repetitive domain, NVIQ was not significantly related to this variable, $F(1, 429) = 2.81$, $p < .094$, $\eta^2 = .007$. However, diagnostic group was significantly related, $F(2, 429) = 93.2$, $p < .001$, $\eta^2 = .303$. This variable explains 30.3 % of the total variance. The post hoc test shows that there is significant difference between all groups.

Running the post hoc analysis separately for gender, shows a slightly different pattern. For both genders there were significant differences between all groups on communication. For the social domain there was no significant difference between LI and TD for boys, however, there was significant differences between all groups for girls. Lastly, on the repetitive domain there was no significant difference between LI and TD for either boys or girls.

Table 17.
ADI-R; algorithm scores for non-verbal participants, adjusted for NVIQ

	ASD group *(N=14)*			LI group *(N=39)*			ANCOVA	
	mean	SE	95% CI	mean	SE	95% CI	*F*	P-H[a]
Communication (NV)	5.1	0.9	3.2-6.9	3.7	0.5	2.6-4.8	0.9[*]	A=L[b]
Social	8.5	1.4	5.8-11.3	4.5	0.8	2.8-6.1	3.2[**]	A<L[c]
Repetitive	3.2	0.5	2.2-4.2	1.4	0.3	0.8-2.0	5.4[***]	A<L[c]

[*] *p* .190
[**] *p* .048
[***] *p* .007
Notes:
[a]P-H: = Bonferroni post hoc test (A=ASD, L=LI)
[b]The ASD group was not significantly different from the LI group
[c]The ASD group were significantly lower than the LI group.

Table 17 shows the ADI-R algorithm scores for non-verbal participants. For the nonverbal participants, an ANCOVA was also run to determine the group differences, controlling for

NVIQ. For the children considered non-verbal, abnormality of communication was not significantly associated with NVIQ. $F(1, 50) = .013$, $p < .911$, $\eta^2 = 000$. Neither was diagnostic grouping, $F(1, 50) = 1{,}8$, $p < .190$, $\eta^2 = .034$, showing no significant difference between ASD and LI, on presence of abnormalities in communication. None of the TD children were rated as non-verbal, hence no comparison of algorithm for presence of atypical language was possible between TD and ASD/LI.

For the social domain, NVIQ was not significantly related, $F(1, 50) = .013$, $p < .911$, $\eta^2 = .000$. However, diagnostic groups was significantly related to presence of abnormalities in social domain, $F(2, 50) = 6.332$, $p < .015$, $\eta^2 = .112$. This explained 11.2 % of the total variance. The post hoc test shows that there are significant differences between all groups.

Lastly on the repetitive domain, NVIQ was not significantly related to the variable, $F(1, 50) = .222$, $p < .639$, $\eta^2 = .004$. However, diagnostic groups was also here significant, $F(1, 50) = 9.730$, $p < .003$, $\eta^2 = .163$. This variable explains 16.3 % of the total variance. The post hoc test shows that there is a significant difference between all groups.

Running the post hoc analysis separately for gender shows no difference between ASD and LI on the communication domain for both genders. Additionally there were no gender differences for the social domain for ASD and LI. However, for the repetitive domain there was a significant difference between ASD and LI for boys, but not for girls. For mean values see appendix B and C.

An additional comparison was done to investigate whether there was overlap between the ASD and LI group. This was done using the ADI-R algorithm cut-off score, where the participants were recoded into groups depending on whether they were above or below cut-off. The cut-off score is described in the ADI-R manual.

Table 18.
ADI-R: algorithm, number and percentages of participants above cut-off for all domains. Divided into verbal and non-verbal participants

	ASD group		LI group		TD group	
	n	%	n	%	n	%
Verbal participants[a]						
Communication domain (V) (8)	27	67.5	24	20.3	3	1.1
Social domain (10)	15	37.5	3	2.5	1	0.4
Repetitive domain (3)	28	70.0	15	12.7	13	4.7
Non-verbal participants[b]						
Communication domain (NV) (7)	5	35.7	8	20.5	-	-
Social domain (10)	6	42.9	5	12.8	-	-
Repetitive domain (3)	8	57.1	9	23.1	-	-

The cut-off score from the ADI-R for each domain are presented in brackets (X)

[a]ASD N=40, LI N=118, TD N=277

[b]ASD N=14, LI N=39

Table 18 shows the numbers and percentages of participants about the cut-off score for autism on the ADI-R algorithm. The ASD group has the highest number of participants scoring over cut-off across both verbal and nonverbal groups of participants. Interestingly, the number of participants scoring over cut-off seems to decrease for the non-verbal participants. For the LI group, there was a number of participants who scored over cut-of for the communication domain, while there were few that scored over cut-off for the social and repetitive domains. Interestingly, the number of participants scoring over the repetitive cut-off increased for the non-verbal participants. This pattern was opposite of the ASD group. While the percentage is small, a number of participants in the TD group scored over cut-off for the repetitive domain.

The last comparison of autism symptomatology compared the number of participants above cut-off for both the ADI-R and ADOS.

Table 19. *ADI-R & ADOS: Number and percentages of participants above cut-off for both instruments. Divided into verbal and non-verbal participants*

	ASD group		LI group		TD group	
	n	%	n	%	n	%
Verbal participants[a]						
Communication domain (8/3)	23	57.5	10	8.5	0	0.0
Social domain (10/4)	12	30.0	0	0.0	0	0.0
Repetitive domain(3/3)	7	17.5	0	0.0	0	0.0
Non-verbal participants[b]						
Communication domain (7/2)	5	35.7	8	20.5	-	-
Social domain (10/4)	6	42.9	1	2.6	-	-
Repetitive domain (3/3)	3	21.4	1	2.6	-	-

The cut-off score from the ADI-R and ADOS for each domain are presented in brackets (ADI cut-off/ADOS cut-off)

[a]ASD N=40, LI N=118, TD N=277

[b]ASD N=14, LI N=39

Table 19 shows number and percentages of participants above the cut-off score on both ADI-R and ADOS. Not surprisingly the ASD group has the highest number of participants above cut-off for both instruments for both verbal and non-verbal participants. However, it is interesting that so few children with ASD are over cut-off for the repetitive domain. Interestingly, for the verbal participants in the LI group 8.5 % scored above cut-off on the communication domain, while 51.3 % of the LI non-verbal participants scored over cut-off. For the social and repetitive domain, none of the verbal participants in the LI group scored above cut-off, but one of the non-verbal participants did.

4.8 Low-functioning ASD sample

The following includes a short presentation of the mean, SD and range for the LF-ASD group. The LF-ASD sample includes 18 boys and 11 girls.

Table 20.
LF-ASD; Sample characteristcs on all applicable variables

	LF-ASD *(N=29)*			
Variable	n	mean	SD	range
Age (months)	29	42.0	3.7	35-52
FISQ	29	50.4	55.3	49-55
NVIQ	29	55.3	1.9	55-65
Age of first word	16	22.6	10.1	12-48
Age of first phrase	8	29.8	10.8	12-44
CDI DQ	18	34.4	10.8	23-66
VABS domains				
Expressive DQ	29	24.5	13.9	3-63
Receptive DQ	29	31.0	12.7	2-55
PPBS factors				
Social play	12	9.6	2.6	7-15
Solitary - passive	19	12.5	4.5	5-20
ADI-R algorithm domains				
Communication - verbal	3	14.0	7.5	6-21
Communication - nonverbal	26	11.1	2.3	6-14
Social	29	16.4	5.0	4-25
Repetitive	29	5.2	5.0	4-25

Table 20 shows the sample characteristics from the LF-ASD group. Especially the mean score for the ADI-R algorithm scores are high. This reflects that these children show more symptoms of autism, and greater impairment compared to the higher functioning ASD group (see table 16). The Vineland DQ and CDI DQ scores are very low, reflecting profound language difficulties. However, this is not so surprising considering the low scores on both IQ measures. When examining the range for FSIQ for this sample it is very narrow (49-55). This

is related to the instrument (Mullen), and it does not measure scores that go below 49. Their true FSIQ score might therefor be even lower than 49.

The PPBS scores were also low, especially the "social play" factor.

Results from the various language factors derived from the CDI shows that on most questions all participants score zero for not mastering the task. A factor worth noting is vocabulary were 2 participants scored on "uses at least five words as names for familiar objects", 5 scored on "uses at least 10 words", 4 on "has a vocabulary of 20 or more words" and 1 on "uses 50 or more different words in everyday conversation". However, none of the participants scored on any of the length of utterance factors.

It is also worth noting that around half of the sample has developed first words, and only eight have developed phrases.

5 Discussion

The following part will include a discussion of the results with a primary focus on the main findings from the current study.

5.1 Main findings

The aim of the current study was to investigate early language characteristics in ASD compared to LI and TD. Additionally, potential overlap of autism symptomatology and social difficulties were investigated, as well as gender differences.

1. The language functions show both differences and similarities between the ASD and LI group. The LI group had less impaired receptive language, compared to the ASD group. However, they were equally impaired on measures of expressive language. The current study found the LI group to be significantly more delayed compared to the ASD group on length of utterance, vocabulary size and grammar. However, for pronouns the ASD group was more delayed than the LI group. For age of first word the LI and ASD groups were equally impaired, and for age of first phrase the LI group was significantly more delayed than the ASD group. The TD group was significantly better on all measures.

2. On social play behavior in pre-school there were significant differences between all groups. The most socially impaired group was children with ASD. The LI group was more social, but was significantly different from the TD group.

3. There were significant differences on autism symptomatology between ASD, LI and TD groups. Naturally the ASD group had the highest scores. However, the LI group scored significantly higher than the TD group on all domains, reflecting a potential overlap between LI and ASD. When running the analysis separately for gender, there were no differences between LI and TD on repetitive behavior.

5.2 Discussion of the main findings in relation to previous research

The following will include a discussion of the results from the current study in relation to previous research.

5.2.1 Are there distinctive language differences between ASD and LI, or are they overlapping?

The first research question in the current study was related to differences in language for LI and ASD, compared to TD. The results show there are significant differences in language competence between the groups, and this will be discussed thematically.

The data from the current study suggest both the ASD and LI group have delayed acquisition of first word and first phrase. There was no significant difference between LI and ASD on age of first word. However, for age of first phrase there were significant differences between all groups. For both measures the LI group was most delayed. For the LI group the mean age for age of first phrase was 9 months later than the TD group. This reflects a great delay in language development. This is similar to what Geurts and Embrechts (2008) found in their study, except they found no significant differences between ASD and SLI for age of first phrase (both had a mean age of about 35 months). They had a sample of older children, and used a wider definition of language impairment, including phonological disorder, dysarthria and stuttering. This might explain the different results. Interestingly, Kristoffersen et al. (2012) found that 50 % of their typically developing sample of children started combining words at 19 months, which is comparable to the results from the current study of TD children.

The current study found that both the LI and ASD groups have shorter sentences at both three and four years of age, when compared to TD. These results could be understood in relation to the theories about mean length utterance of words (Rice et al., 2010), in the sense that it might reflects the child's general language competence. However, caution should be taken when comparing the length in word values across the current measure and values from MLU studies, as they are not measured the same way. MLU is typically calculated as a mean using a 20-30 min speech sample from the child (Rice et al., 2010). This would give a lower numeric value than the current study, which uses parental reports of length of sentences. However, results from the current study show that the lengths of utterances are very different

between the groups. Comparing the length of utterance to results on other language abilities in the current study might show that it reflects the child's overall language abilities.

When comparing vocabulary size, there were significant differences across all groups. It was generally stable up to 50 words for the ASD and LI group for both three and four year olds. However at 50 words there seems to be a decrease in the number of participants that master this skill for both groups. In contrast, the TD group had high percentages at 50 words across both age groups. According to Owens (2012), a 50 word vocabulary is expected for 18 month olds. With this in mind both the ASD and LI group can be considered significantly impaired. This is especially interesting when considering the average two year delay on various language acquisitions found in the Rice (2013) study. A note should be made that she assessed receptive vocabulary while the current study looks at expressive vocabulary. However, there might not be a difference in the pattern of delay. For ASD some studies have shown that early vocabulary skills are predictive for overall language development (Eigsti et al., 2011). Vocabulary skills for ASD have been said to be connected to pre-linguistic development, which is what Smith, Mirenda, and Zaidman-Zait (2007) found in their study. They found that abilities in vocal imitations, pretend play with objects, and the number of gestures were predictive of vocabulary growth. Interestingly, this could be seen in relation to the usage/gestural approach to language development and their understanding of gestures as basis for vocal language development (Bohannon & Bonvillian, 2013).

For grammar, the current study shows there were differences between all groups. Surprisingly the LI group was more delayed than the ASD group. Previous research has shown that past-tense is difficult for children with SLI. The Rice et al. (2000) study found that children with SLI often used the stem of the word (walk for walked). Interestingly, Tager-Flusberg and Joseph (2003) investigated children with ASD regarding past-tense, and found that a portion of the children with ASD only used the the stem of the word in the same way the children with LI does. However, this was only a portion of the ASD sample, while others had non-impaired past test functioning. They hypothesized that there are different language phenotypes in ASD. Perhaps the findings from the current study could point in the same direction, since the LI group has the lowest scores, this could indicates that they have greater difficulties with past tense compared to the ASD group. While the ASD group has a higher mean value, indicating it contains both children with and without difficulties with past tense.

Regarding use of plurals, which in the current study is grouped under grammar, there were differences in percentages across the groups. The LI and ASD groups both had low percentages of correct use of both regular and irregular plurals. Conti-Ramsden and Hesketh (2003) found that plurals had low strength in predicting SLI, and that children with SLI were almost as good as typically developing children. However, in the Kauschke et al. (2011) study of German children with LI found that they performed in line with MLU matched controls, but they were clearly delayed in their acquisition. This shows that they have similar development in acquisition with typical children, but are delayed. Both studies found that when the children made errors they seemed to overuse the –s suffix. The difference between the results from the current study and other findings in the literature might be related to use of parental reports, or there may be language specific differences between Norwegian and English. Kristoffersen et al. (2012) found that 50 % of their typically developing children mastered the skill at 24 months. For the TD group in the current study 90 % mastered the plural skill at 36 months.

Interestingly the ASD group had greater difficulties with pronouns than the LI group in the current study. There also seems to be differences in what type of pronouns that was difficult. Although most of the four year old children with LI and ASD seemed to master "my" and "mine", they had more difficulties with "we" and "you". When comparing these results to the "pronoun reversal" item from ADI-R it seems to correspond. "Pronoun reversals" was present in all groups, although at a higher level in the LI and ASD groups compared to the TD group. Most participants used their name instead of "I" when referring to themselves. Evans and Demuth (2012) found that pronoun reversal is typically associated with ASD, but some TD children have been found to make the same errors. They found that while some TD children had problems with "I", most were related to the use of "you" instead of "me". It would also have been interesting to see if it was possible to compare birth order with amount of pronoun errors, similar to the Oshima-Takane et al. (1996) study were they found differences between first born and later born siblings in pronoun errors.

When comparing narratives, it is interesting that there seems to be big differences between 3 and 4 years on these questions for all groups. In the sample of three year olds very few mastered the skills, while in the sample of four year olds almost everyone mastered it. However, the ASD and LI group performed lower than the TD group on this variable as well. According to Norbury and Bishop (2003) narratives is a known difficulty for both of these

groups. Interestingly they also found great variation in narrative abilities for typical children, so it is a bit puzzling why so many master this skill in the current study. However, this could be related to the questions used from the CDI. They might not be specific enough to capture the complexity of narratives. Additionally, there might be qualitative differences in how the child use of these skills. When telling a story, a child with LI might use fewer words or need help from the parent, while the TD child might tell it without problems. However, the mother with a child with LI might still rate him as telling stories. Additionally there might be a ceiling effect for these questions. In other research, narratives have been collected using speech samples, with narrations from books (Botting, 2002). This was not done in the current study, and might at least explain some of the differences.

The current study found that the ASD group seemed equally delayed on both receptive and expressive language (VABS). While the LI group scored low on expressive language, they were within the lower normal range on receptive language. The ASD group seemed to be more delayed in receptive language compared to the LI group. This is in concurrence with what Loucas et al. (2008) found in their study using direct assessment. Since the VABS measures how the child uses language in everyday situations, it might be a reflection of language use, rather than language competence. However there is good association between scores on VABS expressive and CDI expressive, which is meant to assess more language competence than language use. This could mean that the instruments are measuring the same aspect of language, or that there is little difference between competence and usage related to small children with ASD and LI, at least in terms of how the informants report language functioning.

Examining individual items of atypical language from the ADI-R reveals great variation in scores. For example, the "inappropriate questions or statements" show most impairment for the TD group, perhaps this reflects greater difficulties with knowing what is appropriate to ask in a social setting in the current studies age group. This question might show a developmental phenomenon that is present for typical children at the current age, but is outgrow at a later age. It could also be hypothesized that the lower scores in the ASD and LI group might be related to their generally lower language abilities.

For other questions like "reciprocal conversations" and "current communicative speech" the ASD and LI group were more impaired than the TD group, which might reflect their communication difficulties. The groups show difficulties with engaging in a conversation

with other partners. However, the reason for the difficulties might be similar on a behavioral level, but may be related to different aspects of communication. The typical interpretation of this would be the children with ASD having difficulties with language use (pragmatics), while the LI group would have more difficulties with language form and content (thus wanting to communicate, but not being able). However, when comparing the similarities between the LI and ASD group on other language functions in the current study, it is difficult not to wonder if this is a simplistic interpretation. The interaction between various parts of the language (form, content and use) might be a potential framework in explaining some of overlapping language difficulties between the two groups.

When examining items such as "verbal rituals" and "intonation/volume/rhythm/rate", higher proportions of the ASD group appear impaired. These items might be more specifically identifying atypical language associated with ASD at this age.

Lastly, it was positive to find that there were a bigger portion of the four year olds that mastered the various language functions, compared to the three year olds. This shows that there is a general positive language development for both the LI and ASD children. However this is only a speculation, as the two samples are different.

5.2.2 Social impairment related to both ASD and LI - consequences in pre-school and long-term outcome?

The results from the current study show differences at levels of social play between all groups. All participants in the groups could engage in behavior related to playing alone, while social play were significantly different across all groups. These results are in line with what has been reported in earlier studies (Fujiki et al., 2001; Stanton-Chapman et al., 2007). The results from the Hart et al. (2004) study are particularly similar. They found language impairment to be significantly related to less social behaviors, where higher level of language impairment significantly inhibited the child's abilities to engage socially. However, no significant relationship between degree of language impairment and amount of withdrawn behavior (comparable to the "solitary-passive" scale used in the current study) was found. However, some caution should be taken when interpreting the results from the current study as the mean values did not differ greatly between the LI and TD group, however they did not have overlapping confidence intervals.

Especially the findings of social difficulties in pre-school for LI children were interesting. It is not difficult to imagine the possible long term outcome early social impairment can have for the development of peer relations, friendships and emotional development. Johnson et al. (2010) hypothesized that the variation within their sample showed that having a language impairment in itself did not predict poor outcome, but they suggested that it might change with increased age of the SLI sample. Additionally it is interesting to wonder if there could be a difference in later outcome for the children with LI in the current study, which has been identified at an early age, compared to others identified later, and those not identified at all. Early identification and intervention might have a positive effect on long-term outcome measures.

In general long term outcome for children with ASD in show they have greater need for support. The study by Orsmond et al. (2013) found they were significantly less likely to never see friends, never get called by friends, never get invited to activities and be socially isolated. Interestingly the Russell et al. (2012) study found that children identified at an early age had worse outcomes, however, earlier identification might be related to level of severity, thereby influencing the results of poorer long-term outcome. Potential long-term follow-up of the children in the current study with regard to long-term outcome will be very interesting. However, since the current study has a sample drawn from a non-clinical population, long-term outcome might be different from findings in the literature. The ABC-study may have identified children with both ASD and LI that would not have been identified until later in the regular school and health system.

5.2.3 Is there overlap in autism symptomatology between ASD and LI?

The third research questions presented was to investigate a possible overlap between ASD and LI on autism symptomatology. As previously noted, an overlap between ASD and LI is highly debated in the literature. The current study shows somewhat mixed results regarding this question. The current study found there were significant differences between all the groups for verbal participants, but there were no significant difference between ASD and LI on communication for the non-verbal participants. However, the LI group had higher autism symptomatology scores than the TD group, but not to the same degree as the ASD group. It would be somewhat expected that the LI group had some difficulties with communication and

social interaction that could be similar to ASD. Yet it was more unexpected that the LI group would have scores on the repetitive domain. Interestingly, when examining gender separately, there were no significant difference between LI and TD on the repetitive behavior domain for either gender. These results are similar to the group values Loucas et al. (2008) found in their study, however their values were a bit higher for some of the domains. The reason for this might be different samples.

The differences between ASD and LI were explored further by comparing the number of participants scoring higher than the cut-off score on each domain in ADI-R. This revealed that while most of the participants in the ASD group scored over the cut-off, there were also a number of participants in the LI group that score over cut-off. This perhaps shows a potential overlap between the groups. However, caution should be taken when interpreting the scores for the non-verbal participants as they are very few in the current study (especially the ASD group). The potential overlap was further investigated by comparing which participants that scored over cut-off on both the ADI-R and the ADOS. The results show that a number of both verbal and non-verbal participants in the LI group scored over cut-off on the communication domain. Further, none of the verbal participants scored above cut-off for the social and repetitive domain, while one non-verbal LI participant scored above cut-off for social and repetitive domains on both instruments. Interestingly, very few of the ASD participants scored above cut-off for the repetitive domain. A possible explanation for this might reflect some methodological problems with using the mean score as a cut-off for this domain. However it might also show that, in general, there is less repetitive behavior in children with ASD at the current age.

Combing information from both the ADI-R and ADOS seems to show that there is less chance of a potential overlap between the ASD and LI groups. However, for the communication domain there were a small number of children in the LI group that scored over cut-off on both instruments. Some of the atypical language items from ADI-R could shed some light on the potential difficulties for some of the children with LI on the communication domain. For these variables it was found that children with LI scored on several of the items, however they had lower levels of impairment compared to the ASD group. Additionally, there are a small number of participants in the LI group that score on par with the ASD group on level of severity. Future research on this dataset could investigate whether there is a connection between these variables, or if this is just coincidental. Examining individual items

that make up the social and repetitive domains might also be a good strategy for examining possible differences between the groups.

In conclusion, investigating the potential overlap from the various instruments strengthens the hypothesis that there is little overlap between the groups. However, it should been investigated further.

5.2.4 Summary

In summary, there are many different findings in the current study. On autism symptomology there is variability in the findings. While none of the groups seem to overlap, there seems to be some similarities between the LI and ASD group. This could be interpreted as the LI groups share some autism symptomology with ASD, and that there is a potential continuum between the diagnoses. Maybe the social and communicative difficulties associated with both groups manifest themselves as identical on a behavioral level, but have different origins? Different origins could perhaps explain the different levels of impairment found in the current study when examining individual items on the ADI-R. It might also be that the language difficulties in ASD only explain a portion of their total social and communication difficulties. Lastly it might also be that they are similar at the current age group, but will develop differently into school age, as the Williams et al. (2008) literature review found. Detangling the potential causal relationship between these variables, and possible cofounders is not possible in the current study.

Another question is whether ASD and LI can be co-morbid difficulties. Maybe there is a subsample of the LI group that shows autism symptomatology. It has been hypothesized that these children might have an additional social communication disorder (or pragmatic language impairment). The same might also be true for the ASD group, that only a subgroup show language difficulties, and greater impaired than others in the spectrum (described as ALI). This is what Loucas et al. (2008) described as the double-hit theory, however their study showed there was no significant association between level of language impairment and autism symptomology. Williams et al. (2008) found in their literature review that there was similarities across the groups, but they were qualitative different. It is also an interesting notion that children with primary LI should represent a qualitative different language learner than other children with other diagnosis and co-morbid language difficulties. According to

Tomblin (2011) it should not be surprising that this group of children can overlap with other diagnosis.

5.3 Discussion of additional findings in relation to previous research

The following will include a discussion about the secondary findings from the current study in relation to previous research.

5.3.1 Are there gender differences in between group scores for ASD, LI and TD?

The current study has found a greater proportion of males in both the ASD and LI groups. Both the ASD and LI groups had exactly the same distribution with 78.3 % males and 21.7 % females (this includes the lower functioning sample in the overall ASD %). For the ASD group this is in line with previous findings of about eight to one (Mandy et al., 2012). However, for the LI group the number of girls is substantially lower than previous research (Tomblin et al., 1997). The reason for this difference is unknown, as their study also had a non-clinical sample. However, it might be related to some other methodological differences.

In the current study each variable was analyzed separately for gender. It was found that the between group differences mostly were the same as when the analysis was run on the whole sample (i.e. not split on gender). However, there were some exceptions; for grammar skills there was no significant difference between ASD and LI for girls. Interestingly, the opposite was true for pronouns, where a significant difference was found between ASD and LI. For social play ASD was found not to be different from LI. For autism symptomatology there were also differences between boys and girls. Perhaps the most notable for the verbal sample was that when the repetitive behavior domain was split by gender, there was no difference between LI and TD. The explanation for this might be loss of statistical power. For the nonverbal sample there was no difference between ASD and LI for either gender, except for the repetitive domain were there was a significant difference for boys but not for girls.

It is difficult to interpret these differences. It could be related to true gender differences or perhaps lack of statistical power. There are very few girls with ASD on all variables.

However, when there are significant differences, in spite of low number of participants, it is more likely to reflect a true difference between the groups.

5.3.2 Does the low-functioning ASD group have the same difficulties as the higher-functioning ASD group?

There were 29 children with ASD who scored lower than 55 on the FSIQ and NVIQ cognitive functioning variables. Since there were no other participants in the same IQ group as the LF-ASD group, it was considered doubtful to compare them with participants in the LI group.

There are some methodological challenges when interpreting the IQ scores for this group, as the study by Wetherby et al. (2004) found, many young children with ASD scored the lowest score possible on the early learning composite in Mullen. As a way of differentiating these better they recoded the age equivalent score derived from the manual into a DQ score. The DQ would take into account the child's chronological age and the variability found in the group. However given the uncertainty in using DQ as a measure (see Mervis and Klein-Tasman (2004)), this was not done in the current study. So there might be more variation in the sample than what the current range implies.

As shown in table 20 their scores reflect great difficulties in all areas assessed. It might not come as a surprise that they scored higher on all autism symptomatology variables compared to their higher functioning ASD counterparts. But as the standard deviation shows, there is great variation within the LF-ASD group as well.

On the age of first word and age of first phrase the mean scores show great delay on both variables. Only half the sample has begun using words, and only six have started using phrases. This is interesting since Mayo et al. (2013) found that producing words before 24 months predicted better cognitive abilities, and adaptive skills for children with ASD.

When comparing the expressive and receptive language variables from the VABS and CDI, their mean scores show great language delay for this group. This is comparable to what Maljaars et al. (2012) found in their study, although they compared the LF-ASD group to children with intellectual disability, and found the LF-ASD group to have more language difficulties. Maljaars et al. results are in line with what has been found in other studies about early language in toddlers with ASD (Weismer et al., 2010).

Although the results from the current study show that they in general have low language abilities, there is variation within this group as well, at least on the DQ scores from VABS and CDI. However, when comparing individual items from the CDI there is little variation on the items used in the current study, except for vocabulary where a small number of participants score on some of the variables.

When comparing the LF-ASD group to the higher functioning children with ASD, they have lower score on almost all variables, reflecting greater language difficulties. However there was some variability in language function within this group as well.

5.3.3 How can the IQ difference in the current study be understood?

The sample demographics showed great differences in FSIQ and NVIQ between groups. The ASD group had low performance on both the FSIQ and NVIQ. However the LI group had better performance on NVIQ compared to FSIQ. This was not surprising, since lower performance on FSIQ is probably due to difficulties with language, and hence the verbal subtests of the instrument. Interestingly, when controlling for NVIQ in the various ANCOVA analysis, they explained little of the total variance. The effect was about 5 % across all variables. However, the diagnostic groupings explained much more of the variance. This could be interpreted as the variables in the current study being more related to diagnosis than NVIQ.

In the TD group, both the NVIQ and FSIQ were higher than the average standard score (mean 100) for cognitive instruments such as SB5. This might reflect the selection bias in the MoBa sample, or the potential participation bias in the ABC-study. It might also reflect difficulties with using American norms on a Norwegian sample. Since all the participants in the TD group were administered the SB5 full-scale or abbreviated IQ measure, the high mean score should not be related to the construction of the IQ composite. However, some caution should be taken with the abbreviated IQ measure, as Coolican et al. (2008) found it had a tendency to overestimate the child's abilities.

5.4 General limitations and strengths of the study

The following includes a discussion of limitations and strengths for the current study.

5.4.1 Regarding causality

Shadish et al. (2002) argue that a causal relationship exists if the cause precedes the effect, the cause was related to the effect and we find no other plausible explanation for the effect, except from the defined cause. The current study has a non-experimental cross-sectional design and therefore does not manipulate any variables. Since data is only collected at one point in time, it is impossible to know if cause precedes the effect (Shadish et al., 2002). Since the current study is descriptive in its focus, the relationship between language and various outcomes is not considered the main focus. However, it must be said that the current study does assume some relationship between diagnosis and language outcome, but the direction of the phenomenon is unknown. An additional outcome that was investigated was impaired social play behavior in pre-school. It was hypothesized that it was related to language difficulties. But it is not possible to prove a potential causal relationship between the two. However, the results do show impaired social play behavior for both ASD and LI.

Language impairment is highly related to motor, social and cognitive functions, so detangling the relationship between the factors is not possible in the current study. Controlling for the NVIQ cofounder and matching on age were two measures taken to increase the validity of the current findings, but the casual relationship between them was not investigated. The question remains: Is IQ is the mediator for language, or is it the other way around?

5.4.2 Strengths and limitations

There are several limitations in the current study. None of the language measures used in the current study relies on direct assessment of the child's abilities. However, in an article by Charman (2004) on assessing pre-school children with ASD, he states that direct assessment might not be feasible, due to the low language functioning of these children. However, there is a potential limitation in the use of parental reports, at least on language comprehension. The language used in familiar situations and the non-verbal compensation (e.g. gestures) many families use, make it difficult to assess the child's "true" language comprehension. The current study has mostly used parental reports of expressive language, which in other studies

has yielded more reliable reports than those assessing receptive language (Ring & Fenson, 2000).

Additionally, there is a potential limitation in the use of age equivalents for constructing the VABS DQ and CDI DQ measures. Since the age equivalent is not on an interval scale, in the sense that there is a difference in the four month development between 16 and 20 months, and 36 and 40 months (Mervis & Klein-Tasman, 2004). In the current study the DQ scores has primarily been used for comparison of group differences, not establishing developmental levels for individuals. However, the use of age equivalents must be considered a potential limitation in the interpretation of the results, especially considering the extreme values it generated for the outliers.

Lastly, there have also been discussions in the literature regarding using IQ as a covariate in analysis. In the current study NVIQ was used as a covariate for all ANCOVA analyses. Dennis et al. (2009) argues that IQ is not an appropriate covariate in studies of neurodevelopmental disorders (like ASD and LI). If this is true, then the current study might have produced overcorrected findings in the ANCOVA analyses, and thereby making the groups seem more similar than they actually were. This can be a potential limitation of the current study. Regardless of the discussion for using IQ as a covariate, comparing the children on NVIQ instead of FSIQ is considered positive, since NVIQ is less influenced by the child's language difficulties (Eigsti et al., 2011). Both the ASD and LI group are highly heterogeneous on language difficulties, but the ASD group might show greater variability as it contains groups of children both with and without language impairment. Comparing only the language impaired ASD group with the LI group would be an interesting approach to look at differences and potential overlap. Likewise, it would be interesting to investigate the children with LI with scores above cut-off on both ADI-R and ADOS.

In addition to the limitations of the current study, there are also several strengths. Using a control group, drawn from the same sample as the cases can be considered a strength. The controls are selected from the same hospital as the potential cases. This covers a big portion of Norway as many of the major hospitals participated in MoBa. However, there is always a risk for participation-bias in the control group, since they might not have the same motivation as the group with difficulties (Shadish et al., 2002). It is difficult to say that the results from the current study can be generalized to the population. The characteristics of the participants might be different from those that do not. This potential selection bias in MoBa and ABC has

been discussed in section 3.4.3. The groups are also well matched on age which should limit (but not eliminate) the general development as a cofounder when comparing participants.

5.5 Potential clinical implications

The main aim of the current study has been to investigate group differences on language characteristics for ASD, LI and TD descriptively. It has to a lesser extent focused on the individual differences. Some might say that for speech-language clinicians, who have to treat children on a case-by-case basis, the current study might have limited clinical relevance, since their focus is more on the specifics and individual scores rather than mean differences on a group level. However, Tomblin (2011) makes an excellent point when stating that one must assume that individual differences are largely systematic, and thus it is not by chance that children with LI and ASD are both similar and different from each other. And he goes on to say that this systematic nature is what drives our hopes that we can learn what processes contribute to these differences. Better understanding of different processes might even result in better interventions and ultimately a better outcome for the children each clinician sees.

Another possible implication of the current study is focusing on the LI group. While a lot of attention has gone into early identification of ASD, less attention has been given to children with language impairments. While language is still early in its development at such a young age, the results from the current study shows that they are almost equally as impaired as the ASD group when it comes to language, and they also show signs of social withdrawal. However, the long term outcome for potential social difficulties for this group needs to be investigated.

The results from the current study shows that group differences between ASD and LI, compared to TD, are present from the children are very young. Early identification of children with difficulties opens the possibility for intervention at an early age and perhaps reduces the risk for academic failure and may contribute to strengthen the child's belief in his or her own abilities.

5.6 Future research

Future research should perhaps move away from the general notion of comparing receptive and expressive vocabulary as domains, and to a greater extent investigate more specific language functions within those domains (e.g. mean length of utterance) as this has the potential to show greater variability between the ASD and LI groups.

Additionally, it might be interesting to explore the potential overlap in autism symptomatology between ASD and LI further, by investigating more of individual items from domains from the ADI-R and ADOS, while also including other instruments that assess autism. Perhaps there are distinctive subgroups within the LI group that show some autism symptomatology, or maybe they are distributed randomly across the group. It would also be of interest to use supplementary data from MoBa to investigate the impact of mother's education, socio-economic status and other potential cofounders. These variables could shed even more light on these phenomena. Additionally, it would be interesting to draw upon the longitudinal data from MoBa, to investigate the language abilities of both LI and ASD children at 18 months, as well as their future development after receiving the diagnosis.

6 Conclusion

In conclusion, this study shows the complexity in how to understand early language characteristics in ASD and LI. Results from the current study shows both similarities and differences between the groups on various language abilities. While there were mostly significant differences between the groups, it was somewhat surprising that we did not find bigger numeric differences between the ASD and LI group. It was also interesting that the results from this study were in line with other studies, often using more comprehensive language assessments. Especially, since the current study relied only on parental reports, and used somewhat crude measures of language functioning. While additional direct assessment of language would have strengthened some of the conclusions, this study shows the potential value of parental reports in assessment of language difficulties.

Traditionally, a lot of research in special needs education has focused on the effect of intervention in school/pre-school. The current study did not have intervention as its main aim, but regardless it might have some clinical value for special needs educations, especially for speech-language pathologists, and others who are interested in language abilities of pre-school children with ASD and LI. It is especially interesting that there is such similarity across diagnoses at this early age, and great variation within each diagnostic group. It could be said that receiving a diagnosis is only getting half way there, it is equally important to assess strengths and limitations for each child to help them receive their optimal intervention. It might be valuable for speech language pathologists to look beyond diagnoses when assessing and treating language difficulties in children, with both ASD and LI. The current study makes a small contribution in highlighting this variability, and underscores that each child needs to be met as an individual.

Bibliography

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders: DSM-IV-TR*: American Psychiatric Publishing, Inc.

Backe-Hansen, E. (2012, 10. mai 2012). "Barn". Retrieved 30. oktober, 2012, from http://www.etikkom.no/no/FBIB/Temaer/Forskning-pa-bestemte-grupper/Barn/

Barrett, M., Harris, M., & Chasin, J. (1991). Early lexical development and maternal speech: a comparison of children's initial and subsequent uses of words. *Journal of Child Language, 18*(01), 21-40. doi: doi:10.1017/S0305000900013271

Bishop, D. (1997). *Uncommon understanding: Development and disorders of language comprehension in children*: Psychology Press Hove.

Bishop, D. (2003). Autism and specific language impairment: categorical distinction or continuum. *Novartis Found Symp, 251*, 213-226.

Bishop, D., Clark, B., Conti-Ramsden, G., Norbury, C. F., & Snowling, M. J. (2012). RALLI: An internet campaign for raising awareness of language learning impairments. *Child Language Teaching and Therapy, 28*(3), 259-262. doi: 10.1177/0265659012459467

Bishop, D., & Norbury, C. F. (2002). Exploring the borderlands of autistic disorder and specific language impairment: a study using standardised diagnostic instruments. *Journal of Child Psychology and Psychiatry, 43*(7), 917-929.

Bloom, L., & Lahey, M. (1978). Language Development and Language Disorders.

Bohannon, J. N., & Bonvillian, J. D. (2013). Theoretical approaches to language acquisition. In J. B. Gleason & N. B. Ratner (Eds.), *The development of language*. Boston: Pearson.

Botting, N. (2002). Narrative as a tool for the assessment of linguistic and pragmatic impairments. *Child Language Teaching and Therapy, 18*(1), 1-21.

Botting, N. (2005). Non-verbal cognitive development and language impairment. *Journal of Child Psychology and Psychiatry, 46*(3), 317-326.

Botting, N., & Conti-Ramsden, G. (1999). Pragmatic Language Impairment without Autism: The Children in Question. *Autism, 3*(4), 371-396.

Cameron, A. (2004). Kurtosis. In M. Lewis-Beck, A. Bryman & T. Liao (Eds.), The Sage encyclopedia of social science research methods (pp. 543-544). Retrieved from http://srmo.sagepub.com/view/the-sage-encyclopedia-of-social-science-research-methods/n467.xml.

Chaffee, C. A., Cunningham, C. E., Secord-Gilbert, M., Elbard, H., & Richards, J. (1990). Screening effectiveness of the Minnesota Child Development Inventory expressive and receptive language scales: Sensitivity, specificity, and predictive value. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 2*(1), 80-85.

Charman, T. (2004). Matching preschool children with autism spectrum disorders and comparison children for language ability: methodological challenges. *J Autism Dev Disord, 34*(1), 59-64.

Charman, T., Drew, A., Baird, C., & Baird, G. (2003). Measuring early language development in preschool children with autism spectrum disorder using the MacArthur Communicative Development Inventory (Infant Form). *Journal of Child Language, 30*(1), 213-236.

Conti-Ramsden, G. (2009). The field of language impairment is growing up. *Child Language Teaching and Therapy, 25*(2), 166-168.

Conti-Ramsden, G., Crutchley, A., & Botting, N. (1997). The extent to which psychometric tests differentiate subgroups of children with SLI. *Journal of Speech, Language and Hearing Research, 40*(4), 765.
Conti-Ramsden, G., & Hesketh, A. (2003). Risk markers for SLI: a study of young language-learning children. *International Journal of Language & Communication Disorders, 38*(3), 251-263.
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.
Coolican, J., Bryson, S. E., & Zwaigenbaum, L. (2008). Brief Report: Data on the Stanford–Binet Intelligence Scales in Children with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders, 38*(1), 190-197.
Coplan, R. J., & Rubin, K. H. (1998). Exploring and assessing nonsocial play in the preschool: The development and validation of the preschool play behavior scale. *Social Development, 7*(1), 72-91.
Coyne, I. (2010). Research with Children and Young People: The Issue of Parental (Proxy) Consent. *Children & society, 24*(3), 227-237.
Dennis, M., Francis, D. J., Cirino, P. T., Schachar, R., Barnes, M. A., & Fletcher, J. M. (2009). Why IQ is not a covariate in cognitive studies of neurodevelopmental disorders. *Journal of the International Neuropsychological Society, 15*(03), 331-343.
Eaves, L. C., & Ho, H. H. (2004). The very early identification of autism: Outcome to age 41/2–5. *J Autism Dev Disord, 34*(4), 367-378.
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.
Evans, K. E., & Demuth, K. (2012). Individual differences in pronoun reversal: Evidence from two longitudinal case studies*. *Journal of Child Language, 39*(1), 162.
Fenson, L. (2007). *MacArthur-Bates Communicative Development Inventories: user's guide and technical manual*: Paul H. Brookes Publishing Company.
Field, A. (2009). *Discovering statistics using SPSS*: Sage Publications Limited.
Fricke, S., Bowyer-Crane, C., Haley, A. J., Hulme, C., & Snowling, M. J. (2013). Efficacy of language intervention in the early years. *Journal of Child Psychology and Psychiatry, 54*(3), 280-290.
Fujiki, M., Brinton, B., Isaacson, T., & Summers, C. (2001). Social behaviors of children with language impairment on the playground: A pilot study. *Language, Speech, and Hearing Services in Schools, 32*(2), 101.
Geurts, H. M., & Embrechts, M. (2008). Language profiles in ASD, SLI, and ADHD. *J Autism Dev Disord, 38*(10), 1931-1943.
Harris, M., Barrett, M., Jones, D., & Brookes, S. (1988). Linguistic input and early word meaning. *Journal of Child Language, 15*(01), 77-94.
Hart, K. I., Fujiki, M., Brinton, B., & Hart, C. H. (2004). The relationship between social behavior and severity of language impairment. *Journal of speech, language and hearing research, 47*(3), 647.
Hick, R. F., Joseph, K. L., Conti-Ramsden, G., Serratrice, L., & Faragher, B. (2002). Vocabulary profiles of children with specific language impairment. *Child Language Teaching and Therapy, 18*(2), 165-180.
IBM Corp. (2011). IBM SPSS Statistics for Windows, (Version 20.0). Armonk, NY: IBM Corp.
Ireton, H. (1992). Manual for the Child Development Inventory: Minneapolis, Minn: Behavior Science Systems.

Johnson, C. J., Beitchman, J. H., & Brownlie, E. B. (2010). Twenty-Year Follow-Up of Children With and Without Speech-Language Impairments: Family, Educational, Occupational, and Quality of Life Outcomes. *American Journal of Speech - Language Pathology (Online), 19*(1), 51-65.

Kauschke, C., Kurth, A., & Domahs, U. (2011). Acquisition of German noun plurals in typically developing children and children with specific language impairment. *Child Development Research, 2011*.

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.

Kristoffersen, K. E., Simonsen, H. G., Eiesland, E. A., & Henriksen, L. Y. (2012). Utvikling og variasjon i kommunikative ferdigheter hos barn som lærer norsk en CDI-basert studie.

Kuhl, P. K. (2004). Early language acquisition: cracking the speech code. *Nature reviews neuroscience, 5*(11), 831-843.

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.

Lord, C. (2002). *Autism diagnostic observation schedule: ADOS: Manual*: Western Psychological Services.

Lord, C., & Jones, R. M. (2012). Annual research review: re-thinking the classification of autism spectrum disorders. *J Child Psychol Psychiatry, 53*(5), 490-509. doi: 10.1111/j.1469-7610.2012.02547.x

Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism Diagnostic Interview-Revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders, 24*(5), 659-685.

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., Lopez, K., & Lord, C. (2007). Characterizing communicative development in children referred for autism spectrum disorders using the MacArthur-Bates Communicative Development Inventory (CDI). *Journal of Child Language, 34*(3), 623.

Magnus, P., Irgens, L. M., Haug, K., Nystad, W., Skjærven, R., Stoltenberg, C., & Group, T. M. S. (2006). Cohort profile: The Norwegian Mother and Child Cohort Study (MoBa). *International Journal of Epidemiology, 35*(5), 1146-1150.

Maljaars, J., Noens, I., Scholte, E., & van Berckelaer-Onnes, I. (2012). Language in Low-Functioning Children with Autistic Disorder: Differences Between Receptive and Expressive Skills and Concurrent Predictors of Language. *Journal of Autism and Developmental Disorders, 42*(10), 2181-2191.

Mandell, D. S., Novak, M. M., & Zubritsky, C. D. (2005). Factors Associated With Age of Diagnosis Among Children With Autism Spectrum Disorders. *Pediatrics, 116*(6), 1480-1486.

Mandy, W., Chilvers, R., Chowdhury, U., Salter, G., Seigal, A., & Skuse, D. (2012). Sex differences in autism spectrum disorder: Evidence from a large sample of children and adolescents. *Journal of Autism and Developmental Disorders, 42*(7), 1304-1313.

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.
Mervis, C. B., & Klein-Tasman, B. P. (2004). Methodological issues in group-matching designs: α levels for control variable comparisons and measurement characteristics of control and target variables. *Journal of Autism and Developmental Disorders, 34*(1), 7-17.
Mitchell, S., Brian, J., Zwaigenbaum, L., Roberts, W., Szatmari, P., Smith, I., & Bryson, S. (2006). Early language and communication development of infants later diagnosed with autism spectrum disorder. *J Dev Behav Pediatr, 27*(2 Suppl), S69-78.
Morrow, V., & Richards, M. (1996). The Ethics of Social Research with Children: An Overview1. *Children & society, 10*(2), 90-105.
Mullen, E. (1995). Mullen Scales of Early Learning (AGS Edition). Circle Pines, MN: American Guidance Services: Inc.
Nilsen, R. M. (2013). Unkown Title. *Submitted*.
Nilsen, R. M., Vollset, S. E., Gjessing, H. K., Skjærven, R., Melve, K. K., Schreuder, P., . . . Magnus, P. (2009). Self-selection and bias in a large prospective pregnancy cohort in Norway. *Paediatric and perinatal epidemiology, 23*(6), 597-608.
Norbury, C. F., & Bishop, D. V. (2003). Narrative skills of children with communication impairments. *International Journal of Language & Communication Disorders, 38*(3), 287-313.
Orsmond, G., Shattuck, P., Cooper, B., Sterzing, P., & Anderson, K. (2013). Social Participation Among Young Adults with an Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, 1-10.
Oshima-Takane, Y., Goodz, E., & Derevensky, J. L. (1996). Birth order effects on early language development: do secondborn children learn from overheard speech? *Child Development, 67*(2), 621-634.
Owens, R. E. (2012). Language development: an introduction. (8th ed.).
Pallant, J. (2010). *SPSS survival manual: A step by step guide to data analysis using SPSS*: Open University Press.
Paul, R., Chawarska, K., Cicchetti, D., & Volkmar, F. (2008). Language outcomes of toddlers with autism spectrum disorders: a two year follow-up. *Autism Res, 1*(2), 97-107.
Psychometric Conversion Table. (2003). Psychometric Conversion Table. Retrieved 20.03, 2013, from http://faculty.pepperdine.edu/shimels/Courses/Files/ConvTable.pdf
Rapin, I., & Dunn, M. (2003). Update on the language disorders of individuals on the autistic spectrum. *Brain & development, 25*(3), 166-172.
Redmond, S. M., & Rice, M. L. (1998). The socioemotional behaviors of children with SLI: social adaptation or social deviance? *Journal of speech, language and hearing research, 41*(3), 688.
Rice, M., Smolik, F., Perpich, D., Thompson, T., Rytting, N., & Blossom, M. (2010). Mean length of utterance levels in 6-month intervals for children 3 to 9 years with and without language impairments. *Journal of Speech, Language and Hearing Research, 53*(2), 333.
Rice, M., Wexler, K., Marquis, J., & Hershberger, S. (2000). Acquisition of irregular past tense by children with specific language impairment. *Journal of Speech, Language and Hearing Research, 43*(5), 1126.
Rice, M. L. (2013). Language growth and genetics of specific language impairment*. *International Journal of Speech-Language Pathology, 15*(3), 223-233.
Rice, M. L., Redmond, S. M., & Hoffman, L. (2006). Mean length of utterance in children with specific language impairment and in younger control children shows concurrent

validity and stable and parallel growth trajectories. *Journal of Speech, Language and Hearing Research, 49*(4), 793.
Ring, E. D., & Fenson, L. (2000). The correspondence between parent report and child performance for receptive and expressive vocabulary beyond infancy. *First Language, 20*(59), 141-159.
Roid, G. (2003). Stanford-Binet Intelligence Scales (; SB5). *Rolling Meadows, IL: Riverside*.
Russell, G., Golding, J., Norwich, B., Emond, A., Ford, T., & Steer, C. (2012). Social and behavioural outcomes in children diagnosed with autism spectrum disorders: a longitudinal cohort study. *Journal of Child Psychology and Psychiatry, 53*(7), 735-744.
Rutter, M., Bailey, A., & Lord, C. (2003). *The Social Communication Questionnaire: Manual*: Western Psychological Services.
Rutter, M., Le Couteur, A., & Lord, C. (2003). *ADI-R: Autism Diagnostic Interview--Revised: Manual*: Western Psychological Services.
Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological methods, 7*(2), 147-177.
Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). Experimental and quasi-experimental designs for generalized causal inference.
Smith, V., Mirenda, P., & Zaidman-Zait, A. (2007). Predictors of expressive vocabulary growth in children with autism. *Journal of Speech, Language and Hearing Research, 50*(1), 149.
Snowling, M., Bishop, D., & Stothard, S. E. (2003). Is preschool language impairment a risk factor for dyslexia in adolescence? *Journal of Child Psychology and Psychiatry, 41*(5), 587-600.
Snowling, M., Bishop, D., Stothard, S. E., Chipchase, B., & Kaplan, C. (2006). Psychosocial outcomes at 15 years of children with a preschool history of speech-language impairment. *Journal of Child Psychology and Psychiatry, 47*(8), 759-765.
Sparrow, S. S., Balla, D. A., & Cicchetti, D. V. (1984). *Vineland Adaptive Behavior Scales*.
Sponheim, E. (1996). Changing criteria of autistic disorders: a comparison of the ICD-10 research criteria and DSM-IV with DSM-III-R, CARS, and ABC. *Journal of Autism and Developmental Disorders, 26*(5), 513-525.
St Clair, M. C., Pickles, A., Durkin, K., & Conti-Ramsden, G. (2011). A longitudinal study of behavioral, emotional and social difficulties in individuals with a history of specific language impairment (SLI). *Journal of Communication Disorders, 44*(2), 186-199.
Stanton-Chapman, T. L., Justice, L. M., Skibbe, L. E., & Grant, S. L. (2007). Social and behavioral characteristics of preschoolers with specific language impairment. *Topics in Early Childhood Special Education, 27*(2), 98-109.
Stoltenberg, C., Schjolberg, S., Bresnahan, M., Hornig, M., Hirtz, D., Dahl, C., . . . Group., A.-S. (2010). The Autism Birth Cohort: a paradigm for gene-environment-timing research. *Mol Psychiatry, 15*(7), 676-680. doi: 10.1038/mp.2009.143
Tadevosyan-Leyfer, O., Dowd, M., Mankoski, R., Winklosky, B., Putnam, S., McGrath, L., . . . Folstein, S. E. (2003). A Principal Components Analysis of the Autism Diagnostic Interview-Revised. *J. AM. ACAD. CHILD ADOLESC. PSYCHIATRY, 42*(7), 865.
Tager-Flusberg, H., & Joseph, R. M. (2003). Identifying neurocognitive phenotypes in autism. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 358*(1430), 303-314.
Tager-Flusberg, H., Paul, R., & Lord, C. (2005). Language and communication in autism. *Handbook of autism and pervasive developmental disorders, 1*, 335-364.
Tager-Flusberg, H., Rogers, S., Cooper, J., Landa, R., Lord, C., Paul, R., . . . Yoder, P. (2009). Defining spoken language benchmarks and selecting measures of expressive

language development for young children with autism spectrum disorders. *Journal of Speech, Language and Hearing Research, 52*(3), 643.
Tomblin, B. (2011). Co-morbidity of autism and SLI: kinds, kin and complexity. *International Journal of Language & Communication Disorders, 46*(2), 127-137.
Tomblin, J. B., Records, N. L., Buckwalter, P., Zhang, X., Smith, E., & O'Brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *Journal of Speech, Language and Hearing Research, 40*(6), 1245.
Unpublished ABC-study protocol. (2010). Unpublished Study Protocol.
Volden, J., Smith, I. M., Szatmari, P., Bryson, S., Fombonne, E., Mirenda, P., . . . Zwaigenbaum, L. (2011). Using the preschool language scale, to characterize language in preschoolers with autism spectrum disorders. *American Journal of Speech-Language Pathology, 20*(3), 200.
Wang, M., Lekhal, R., Aarø, L., & Schjølberg, S. (2012). Co-occurring development of early childhood communication and motor skills: results from a population-based longitudinal study. *Child: Care, Health and Development*.
Webster, R. I., & Shevell, M. I. (2004). Topical Review: Neurobiology of Specific Language Impairment. *Journal of Child Neurology, 19*(7), 471-481.
Weismer, E. S., Lord, C., & Esler, A. (2010). Early language patterns of toddlers on the autism spectrum compared to toddlers with developmental delay. *J Autism Dev Disord, 40*(10), 1259-1273. doi: 10.1007/s10803-010-0983-1
Wetherby, A. M., Woods, J., Allen, L., Cleary, J., Dickinson, H., & Lord, C. (2004). Early indicators of autism spectrum disorders in the second year of life. *Journal of Autism and Developmental Disorders, 34*(5), 473-493.
Williams, D., Botting, N., & Boucher, J. (2008). Language in autism and specific language impairment: Where are the links? *Psychological Bulletin, 134*(6), 944.
World Health Organization. (1992). *The ICD-10 classification of mental and behavioural disorders: clinical descriptions and diagnostic guidelines*: World Health Organization.

Appendix

Appendix A

Sample demographics boys

	ASD group (N=47)			LI group (N=126)			Typical group (N=217)		
	mean	SD	95% CI	mean	SD	95% CI	mean	SD	95% CI
Age (months)	42.3	4.4	41.0-43.6	41.1	2.3	40.7-41.6	42.2	2.4	41.9-42.5
FSIQ	84.4	17.2	79.4-89.4	87.1	11.3	85.1-89.1	103.9	11.5	102.3-105.4
NVIQ	88.1	16.8	83.2-93.1	94.9	11.3	92.9-96.9	106.5	11.6	104.9-108.0

Sample demographics girls

	ASD group (N=7)			LI group (N=35)			Typical group (N=197)		
	mean	SD	95% CI	mean	SD	95% CI	mean	SD	95% CI
Age (months)	44.5	4.5	40.4-48.6	41.0	1.9	40.4-41.7	42.4	2.7	42.0-42.8
FSIQ	78.4	20.4	59.6-97.3	87.2	13.3	82.6-91.8	106.6	11.1	105.0-108.2
NVIQ	77.4	18.1	60.6-94.2	92.1	12.0	88.0-96.2	109.3	10.2	107.9-110.7

Appendix B

Sample variables, run separately on boys

	ASD		LI		TD	
Variable	mean	SE/SD	mean	SE/SD	mean	SE/SD
Age of first word*	16.0	6.6	19.9	7.4	14.9	4.4
Age of first phrase*	25.8	8.0	29.9	6.3	20.9	4.7
CDI						
CDI DQ	69.5	2.8	65.5	1.8	96.7	1.5
Lenght of utterance	1.4	0.1	1.0	0.1	2.6	0.1
Vocabulary	3.4	0.1	2.9	0.1	3.8	0.1
Grammar	1.1	0.1	0.6	0.1	1.9	0.1
Pronouns	1.9	0.2	2.0	0.1	3.4	0.1
VABS domains						
Expressive DQ	71.7	3.3	66.0	1.9	99.1	1.5
Receptive DQ	77.3	3.2	86.7	1.8	97.4	1.5
PPBS factors						
Social play	16.3	1.1	20.7	0.7	24.5	0.4
Solitary - passive	16.5	0.7	15.4	0.4	15.3	0.3
ADI-R algorithm domains						
Communication - verbal	9.1	0.6	4.7	0.3	1.9	0.3
Social - verbal	9.0	0.5	2.2	0.3	1.5	0.3
Repetitive - verbal	4.7	0.3	1.2	0.2	0.7	0.2
Communication - nonverbal	5.0	1.0	4.0	0.6	-	-
Social -nonverbal	8.5	1.5	5.0	0.9	-	-
Repetitive - nonverbal	3.4	0.6	1.5	0.3	-	-

Notes. * standard deviation (SD) instead of standard error (SE)

Appendix C

Sample variables, run separately on girls

	ASD		LI		TD	
Variable	mean	SD	mean	SD	mean	SD
Age of first word*	21.7	11.0	17.9	6.6	14.2	4.2
Age of first phrase*	24.5	8.4	28.8	6.2	19.7	4.4
CDI						
CDI DQ	71.8	11.2	72.9	4.6	102.6	1.9
Lenght of utterance	1.3	0.3	1.1	0.1	2.7	0.1
Vocabulary	3.3	0.3	3.5	0.1	3.8	0.1
Grammar	0.8	0.4	0.7	0.2	2.0	0.7
Pronouns	1.3	0.4	2.5	0.2	3.4	0.1
VABS domains						
Expressive DQ	73.0	8.6	73.1	3.9	104.8	1.6
Receptive DQ	65.3	7.7	92.9	3.4	101.3	1.4
PPBS factors						
Social play	17.3	1.9	21.9	1.0	25.3	0.4
Solitary - passive	15.9	1.6	13.7	0.7	15.5	0.3
ADI-R algorithm domains						
Communication - verbal	11.3	0.8	3.7	0.4	1.1	0.2
Social - verbal	9.7	0.9	2.4	0.4	0.8	0.2
Repetitive - verbal	5.5	0.5	0.8	0.2	0.4	0.1
Communication - nonverbal	4.6	1.4	2.1	0.7	-	-
Social - nonverbal	7.0	1.7	2.3	0.9	-	-
Repetitive - nonverbal	1.5	1.1	1.0	0.6	-	-

Notes. * standard deviation (SD) instead of standard error (SE)

Appendix D

The algorithm ADI-R for "2:0 to 3:11" and "4 or more" in **bold**:
A: Qualitative abnormalities in reciprocal social interaction:

- A1: Failure to use nonverbal behaviors to regulate social interaction
    - Direct gaze
    - Social smiling
    - Range of facial expressions used to communicate
- A2: Failure to develop peer relationships
    - **Imaginative play with peers**
    - Interest in children
    - Response to approaches of other children
    - **Group play with peers OR Friendships (depending on age of subject)**
- A3: Lack of shred enjoyment
    - Showing and directing attention
    - Offering to share
    - Seeking to share enjoyment with others
- A4: Lack of socioemotional reciprocity
    - Use of others body to communicate
    - Offering comfort
    - Quality of social overtures
    - Inappropriate facial expressions
    - Appropriateness of social responses

B: Qualitative abnormalities in communication

- B1: Lack of, or delay in, spoken language and failure to compensate through gestures
    - Pointing to express interests
    - Nodding
    - Head shaking
    - Conventional/instrumental gestures
- B4: Lack of varied spontaneous make-believe of social imitative play
    - Spontaneous imitation if actions
    - Imaginative play

Imitative social play Verbal subjects only:

- B2(V): Relative failure to initiate or sustain conventional interchange
    - Social verbalization/chat
    - Reciprocal conversation
- B3(V): Stereotyped, repetitive, or idiosyncratic speech
    - Stereotyped utterances and delayed echolalia
    - Inappropriate questions of statements
    - Pronominal reversal
    - Neologisms/idiosyncratic language

C: Restricted, repetitive, and stereotyped patters or behavior

- C1: Encompassing preoccupation or circumscribed pattern of interest
    - Unusual preoccupations
    - Circumscribed interests
- C2: Apparently compulsive adherence to nonfunctional routines or rituals
    - Verbal rituals (if verbal)
    - Compulsions/rituals
- C3: Stereotyped and repetitive motor mannerisms
    - Hand and finger mannerisms OR other complex mannerisms or stereotyped body movements (the highest of the two scores)
- C4: Preoccupations with part of objects or nonfunctional elements of material
    - Repetitive use of objects or interests in parts of objects OR unusual sensory interests (the highest of the two scores)

Appendix E

Norwegian version of the CDI – expressive scale

1. Kaller voksne i barnehagen/park (eventuelt dagmamma) med navn
2. Snakker i lange setninger for å uttrykke tanker - minst 6 ord lange setninger
3. Forteller en kort historie som for eksempel Rødhette og ulven; forteller hendelsen i riktig rekkefølge og hvordan historien ender
4. Snakker om fortiden med riktig endinger og bøyning; for eksempel "Jeg lekte med Tor", "Jeg gjorde det", "Jeg gikk til..."
5. Bruker ordet "du" i setningene
6. Beskriver ting helt i detalj, for eksempel "Dukken har hår og en kjole", "Hunden har en hale" etc.
7. Bruker ordene "Kan ikke", "Gjør ikke", "Vil ikke"
8. Sier to eller flere ord ved siden av å si "mamma" og "pappa"
9. Bruker ordene "en", "et", "den", for eksempel "Se en hund", "Se den katten"
10. Bruker minst 10 ord
11. Kommer med ytringer som "Hvis jeg gjør det… så kan jeg…" eller "Når jeg skal ..…. så "
12. Prater; lager lyder som om han/hun snakker i setninger, eller snakker med faktiske setninger
13. Bruker ordene "meg", "min" og "jeg" riktig
14. Bruker flertallsending korrekt; for eksempel "menn" og ikke "manner", "mus" ikke "muser", "føtter" og ikke "foter"
15. Snakker om ting som "kunne skje" eller "kanskje skje", for eksempel "Han kunne skadet seg hvis han ikke var forsiktig"
16. Forteller hva som skjer av aktivitet på et bilde, for eksempel "Kattepusen spiser"
17. Synger enkle sanger
18. Bruker ordet "ikke" i setningene
19. Uttrykker lett tanker og ideer i fullstendige setninger, bruker god grammatikk og uttaler de fleste ord helt tydelig
20. Stiller spørsmål som begynner med "Hva" eller "Hvor"
21. Snakker i setninger som er på minst fire ord
22. Begrunner ting ved å bruke ord som "Fordi…."
23. Snakker helt klart og tydelig; ordene blir forstått i de fl este tilfellene
24. Bruker minst fem ulike ord som er navn på ting (objektord)
25. Bruker minst ett av de følgende ordene: "Meg", "Jeg", "Han", "Hun", "Du" eller "Det"
26. Stiller spørsmål som begynner med "Hvorfor", "Når" eller "Hvordan"
27. Har et ordforråd på 20 ord eller mer
28. Snakker i lange, komplekse setninger, 10 ord eller mer i setningen
29. Snakker i detalj om ting som har skjedd, beskriver en serie av hendelser, f.eks. "Vi gikk til… og så ….Etter det fikk vi…."

30. Snakker om sine egne ting som “Mitt” eller ”Mine”
31. Bruker flertallspronomen som “Vi”, “De”, “Dem” eller “Oss” korrekt
32. Bruker 50 ulike ord eller flere i daglig samtale
33. Kan hviske (snakke lavt)
34. Bruker navnet på figurer som sirkel, firkant, trekant og stjerne
35. Stiller enkle spørsmål og bruker da helt riktige setninger (grammatikk)
36. Peker på ting
37. Spør etter hva ord betyr
38. Bruker flertallsending og legger til “-er”, f.eks. “hest**er**”, ”bil**er**”
39. Gjengir et barnerim som f.eks. “Mikkel rev”
40. Spør etter “Mer” eller “En til”
41. Snakker i setninger med ordene i riktig rekkefølge
42. Ber om noe å drikke eller spise, ved å bruke ord eller lyder
43. Snakker i to- eller tre- ords setninger. (..eller i lengre setninger)
44. Kan si navnet på kjente ting i en billedbok (trenger bare være noen få)
45. Sier “Kan jeg få…” eller “Takk”
46. Kan navnet på minst fem kroppsdeler, f. eks øye, nese, munn, hender eller føtter dersom en spør
47. Binder sammen to setninger med ordene “og”, “eller” eller “men”
48. Har et stort ordforråd som er mer enn bare å kunne telle
49. Sier (uttaler) de fl este ordene han/hun bruker helt korrekt
50. Sier navnet på ukedagene i riktig rekkefølge