8-27-2013

Social and Linguistic Factors in the Development of Children with Autism

Emma C. Kelty-Stephen
University of Connecticut-Storrs, kelty@grinnell.edu

Social and Linguistic Factors in the Development of Children with Autism

Emma C. Kelty-Stephen

University of Connecticut, 2013

There is an ongoing debate in the literature over the main source of information that children use when acquiring and developing language. Theories either support a computational linguistic perspective – in which children are thought to use aspects of language itself and their ability to perceive patterns to acquire language – or a social perspective, in which children are thought to use their ability to jointly attend with communication partners in order to acquire language. It is difficult to tease apart these sources of information in typically-developing children, but children with autism spectrum disorders have difficulty with both joint attention and language development, to varying degrees. The current study is a longitudinal analysis of specific language development in young children with autism and their typically-developing peers, using joint attention behaviors and computational behaviors to predict language growth. Joint attention proved to be important in lexical development but not as much in grammatical development, and computational abilities proved important specifically for pronoun use. Furthermore, the growth itself of early grammatical and lexical abilities predicted scores on later language tests. These results support a role for social abilities in the growth of lexical development, and a role for computational abilities on initial language abilities, but not on their growth over time.

Social and Linguistic Factors in the Development of Children with Autism

Emma C. Kelty-Stephen

B.A., Grinnell College (2004)

M.A., University of Connecticut (2011)

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree Doctor of Philosophy

at the

University of Connecticut

2013

APPROVAL PAGE

Doctor of Philosophy Dissertation

Social and Linguistic Factors in the Development of Children with Autism

Presented by

Emma C. Kelty-Stephen

Major Advisor ____

Letitia Naigles

Associate Advisor ____

Marie Coppola

Associate Advisor ____

James Dixon

Associate Advisor ____

Inge-Marie Eigsti

University of Connecticut

2013

Acknowledgements

This work could not have been done without the parent and child participants who volunteered for our study. I am grateful for their time and energy and willingness to let our research team bring an entire lab's worth of equipment into their homes every four months.

My advisor Letitia Naigles has shown faith in me from the very beginning of my graduate studies, and has led me further than I ever expected toward a successful academic career. Our weekly meetings with Marie Coppola and Heather Bortfeld and their lab members have been invaluable to my research development, and Inge-Marie Eigsti and Whit Tabor offered suggestions that expanded my reading and theoretical knowledge beyond developmental work.

Finally, of course, I would like to thank my family and friends. My parents never waiver in their support of me, and that has been crucial throughout this process. My friends Lauren Broder, Jason Anastas, and Anthony Goodwin have been with me every step of the way, and I am very grateful for that. Finally, Damian deserves a second Ph.D., one in patience and support, for everything he has done over the last five years.

Table of Contents

1. Introduction

Language is a symbol system used to communicate ideas, information, and thoughts. In order to become communicating members of a society, children must learn the phonological, lexical, and grammatical, and pragmatic elements of their language. In other words, acquiring a language involves parsing a stream of sounds or signs to find discrete units of meaning, discovering the referents of those units, and figuring out how to combine those units in a way that delivers a desired message in a socially meaningful way. All this learning takes place relatively quickly and successfully in a typically-developing child, who goes from reacting to the prosody of voices (Cooper & Aslin, 1990) to communicating through noises to single words in about one year, and to combining those words in about two years (Bloom, 1998). During toddlerhood and the early school years, typically-developing children increase their vocabularies dramatically (Anglin, 1993), and gain impressive pragmatic and narrative abilities (Berman, 2009). However, certain groups of children seem to lack some of the information or processing abilities to efficiently acquire and develop language. Children with autism spectrum disorders (ASD) begin producing language later than typical children, and show a complicated pattern of language abilities and disabilities (Gerenser, 2009; Eigsti, Bennetto, & Dadlani, 2007). The goal of this dissertation is to examine the development of specific language abilities over time in both typically-developing children and those with ASD, to see what sources of information seem to predict the growth over time of those language abilities, and to further investigate whether the pattern of development can be used to predict later language and cognitive levels.

The early information and abilities that children use to acquire language are a source of great debate in the field of language research. Where do children get the information necessary to accomplish the feat of language acquisition? Possibly from abstract concepts that the child has access to, thanks to the nature of the human mind and the properties of the language system (Pinker, 1994; Lidz, 2007, Crain & Lillo-Martin, 1999), or possibly through social abilities that come into play when children and adults are sharing contexts that allow for mutual understanding of meaning (Tomasello, 1992a; Baldwin & Meyer, 2007). The mechanisms of language acquisition need to be efficient and available from a young age in order to allow for the development of the vocabulary and grammatical rules that allow a child to become proficient in language. Children who acquire language atypically play an important role in this field of research, both clinically work and empirically (Naigles & Bavin, 2013). Research on children who acquire language atypically can help clinicians develop ways of treating and counseling those children, while language acquisition theorists must develop theories that account for situations that deviate from the norm. Language development proceeds in a variety of ways, and theories that account for only a portion of children's behaviors and experiences are unsatisfying. Furthermore, atypical development of language can point to ways that various abilities (such as lexical and semantic abilities) may be related or unrelated, and disorders that affect one source of information in development (e.g., social interaction) more than others can identify the importance of specific abilities in development. The study of atypical language development, then, can illuminate the relative importance of specific abilities for different areas of language development. This dissertation instantiates this goal by investigating the relative importance of social abilities for lexical and grammatical development in TD children and children with ASD; in this latter group, social deficits are well attested (Happé & Frith, 2006).

Theories of language development

The question of how children begin acquiring language has given rise to two major schools of thought, one of which defines children's linguistic or computational abilities as the main contributor to language ability, and the other of which cites social factors as the driving force behind acquisition. Hirsh-Pasek and Golinkoff (1996) coined the descriptive phrases "inside-out" and "outside-in" to refer to these theories, referring to the idea that the knowledge and processing ability comes from inside the child or from the child's environment (i.e., outside the child). Inside-out theories posit that, by virtue of being human, children have some knowledge and abilities that are specific to learning language ("inside" them), and that these abilities are prompted by some type of input or stimulation from the environment, while outside-in theories claim that the information children need to acquire language is in the environment. Another way of thinking about this is to say that inside-out theories of language acquisition place emphasis on children's *computational* and data-collecting abilities, such as their ability to narrow down the possible meanings of a new word based on previous experiences, or their ability to calculate the likelihood of a new word belonging to a certain category. Meanwhile, outside-in theories of language development emphasize the *social* capacities of children, and their ability to deduce the meaning of a new word based on a shared communicative and cultural context.

I will use the terms "social" and "computational" to refer to these groups of theories, rather than "nativist" and "empiricist", which are more commonly used, because the former are more flexible. Nativism is often used to refer to theories that claim innate language-specific structures are present in humans that allow for somewhat automatic language learning, while empiricism is used to refer to theories that claim no innateness whatsoever. However, these theoretical viewpoints do not have to be diametrically opposed. Some theories that posit innate structures (i.e. could be called "nativist") focus more on general learning mechanisms and biases, and in the case of the theories discussed below, it is not at all necessary to think of these mechanisms that are "inside" the child to have been automatically or innately endowed. It is only necessary to consider that the mechanisms are being used *by* the child to process the linguistic information being perceived in the world (Hoff, 2006). Additionally, most empiricists claim some domain-general innate abilities and tendencies towards social contact and association-building (Tomasello, 1992a). Rather than focus on *where* the abilities originate, I am interested in what role these two types of information play in language acquisition and development. In this discussion, I will focus on outside-in theories that consider social mechanisms to be the most important "outside" mechanism; however, it is true that linguistic input and other environmental features may be important as well. The following will outline some evidence for the importance of each of these types of abilities.

Social factors in language development

Social theorists have examined data about the information available in the social environment and proposed that it is sufficient for the child learner to acquire language without preexisting linguistic knowledge and by using domain-general learning mechanisms. On this account, the main challenges that a child faces are reading another's communicative intentions, and recognizing associations between language use and context in their input (Tomasello, 2006; Akhtar, 2004) by using skills such as eye contact, gaze following, and referential pointing. When a child begins to use language, it is to serve a purpose: to communicate and interact socially with others. While computational theories often refer to linguistic input as being "impoverished" relative to what children need to learn language, social theorists do not believe there is any "poverty" of the language stimulus to account for as long as children are sharing communicative contexts and have a structured cultural environment to depend on for clues to meaning (Tomasello, 1992a).

Social theorists claim that sharing communicative acts is the beginning of language development, and that children are trying to get a job done (i.e. communicate) rather than acquiring a skill that has any unique processes or basis (Akhtar & Tomasello, 2000). Rather than bringing prior linguistic information to the task, the child must engage in joint attention and intention-reading in order to fulfill his/her communicative needs. In brief, children do not learn words and then put them into sentences, rather they learn utterances (defined as the "smallest unit in which a person expresses a complete communicative intention" (Tomasello, 2009, p. 72)) and then extract words from those utterances. Thus the processes of learning lexical items and grammar are not separate; they are both achieved by sharing contexts and intentions with other speakers, and picking up on associations between contexts and utterances.

Typically-developing children show excellent joint attention skills before two years of age (Mundy, Block, Delgado, Pomares, Van Hecke, & Parlade, 2007); these skills involve being able to follow another person's eye gaze or point to share an object of attention, or to draw another's attention to an object by pointing or speaking (Carpenter, Nagell, Tomasello, Butterworth, & Moore, 1998; Liszkowski, Carpenter, Henning, Striano, & Tomasello, 2004). Joint attention is thought to be crucial in sharing the communicative contexts needed to acquire language using social abilities. The ability to share in joint attention and share common reference points is critical to acquiring words, and many studies have demonstrated a link between early joint attention and later language abilities. Akhtar, Dunham, and Dunham (1991) found that children whose mothers followed more of the child's attentional bids at age one had higher vocabularies nine to ten months later. Carpenter and colleagues (1998) found a predictive relationship between the overall amount of time spent in joint attention episodes and later linguistic and gestural production. Studies by Baldwin and colleagues (Baldwin, 1993; Baldwin, Markman, Bill, Desjardins, Irwin, & Tidball, 1996) have demonstrated children's ability to understand an adult's referential intent by showing that children assign a novel word to an object that the adult is looking at, even if the child is looking elsewhere or cannot see the object. These results show the importance of social abilities as cues to word learning in particular.

A study of one child's early verb use (Tomasello, 1992b) demonstrated that the child used fairly conservative extensions of new verbs, meaning that she mainly used verbs in attested contexts and was slow to extend those verbs to new usages. On this account, the route to basic grammar learning (i.e., learning to put words together) is based on specific constructions first, rather than abstract rules. Gradually, more experiences with the verb allow for expanded use, and a child produces phrases that are based around the verbs they have learned, but the flexibility of the constructions is limited (see Tomasello, 2000 for review). Social theorists posit that the process of acquiring language is one of mastering utterances (words and phrases) via shared intentions with communication partners, and then putting them together into increasingly complex structures (syntax), rather than working towards a formal system that is mathematically elegant and highly abstract.

Computational factors in language development

Many computational theories are based around the speed and accuracy with which children acquire grammatical rules and lexical items, especially as compared to the quality and quantity of input they hear. In the most extreme case, Chomsky (1981) claimed that acquisition of grammar was "automatic", given a set of universals (Principles) that all children have, and then input sets the rules unique to that particular language's grammar (Parameters). Language structure is considered to be abstract and rule-based, given the patterns of syntactic structure across different languages (Lidz, 2007).

Other theorists have assigned a special role for grammar rules (separate from lexical learning) based on evidence that the input provided to children is nowhere close to providing the information they need to learn language as quickly and accurately as they do (the "poverty of the stimulus" theory: Chomsky, 1959; Crain & Lillo-Martin, 1999). Marcus (1993) demonstrated that children do not receive reliable negative evidence about what makes an ungrammatical statement. For instance, children occasionally make grammatical errors in their language production, such as "maked" instead of "made" (Brown, 1973), but according to Marcus' analysis they almost never hear their parents explicitly correct or negate those grammatically incorrect usages. However, children do not continue making such errors, and in fact they are quite rare in the first place. On this account, then, children cannot rely on their input to determine correct grammatical structure, meaning that they must use additional computational abilities to discover the rules involved in grammar. Snyder (2007) also pointed out that children's earliest uses of grammatical constructions are rule-based, and posited that children show very few errors because they wait until they have abstracted a rule before using a given construction. These pieces of evidence support the idea that children exploit internal computational processes such as abstract rule learning to acquire grammar.

Another computational process, one that need not be domain-specific to language, is learning of patterns via statistical learning, or use of distributional and conditional information in the input (Theissen, 2009). Saffran, Aslin, and Newport (1996) demonstrated that infants at only eight months of age could learn distributional patterns in a nonsense string of syllables very quickly such that they showed a preference for triads of syllables they had heard before. The infants were exposed to a stream of nonsense syllables with carefully manipulated patterns of co-occurrence. During a test phase, infants showed longer listening times to those syllable triads that were consistent with the triads in the stream. These results suggest that infants can parse a stream of speech with no apparent gaps into something analogous to words by “calculating” the probability that one sound will follow another. Gomez and Gerken (1999) expanded this to demonstrate that one-year-old children extracted conditional patterns analogous to grammatical rules in natural languages. Using nonsense syllables, they created conditional probabilities in which certain sounds co-occurred, but with differing sounds in between the target sounds. For instance, sound “A” and sound “B” might occur with sounds “C”, “D”, or “E” in between them; infants recognize the “A-B” pattern even when tested on the pattern “A X B”. The one-year-olds recognized these patterns after brief exposure to the “speech” streams. This demonstrates that children may make use of grammatical abilities independent from lexical items, contrary to the idea from social theorists that children begin language acquisition with utterances and then extract words. The statistical learning literature demonstrates that young children can learn grammatical rules devoid of any communicative context.

Other computational theories are more focused on lexical learning, and posit a variety of biases and expectations that a child is equipped with in order to simplify the task of mapping a word to a referent, which is thought to be too ambiguous a task for children to accomplish given only the information from the environment (Gleitman, 1990). No one supposes that children have innate or natural abilities to come up with the lexical items of a particular language (why else would a Spanish baby say “gato” while an English baby says “cat”) so those must be learned from input; the social context, however, may not be enough to provide the child with the link between sound and meaning necessary for the child to learn lexical items with the speed and accuracy that most do. Hoff and Naigles (2002) looked specifically at toddlers interacting with their mothers in naturalistic contexts with the potential for word learning. They found certain aspects of input to be important to vocabulary growth, but no systematic effect of social experience, suggesting that an internal process is at work. Others have suggested that children have biases that limit the hypotheses children must test when concluding a word’s meaning, such as conceptual biases (Waxman, 2004) or perceptual biases (Smith, 2000), or that they have finely-tuned ways of using grammatical structure to correctly infer the meaning of new words (Naigles, 1990; Naigles & Swensen, 2007; Yuan & Fisher, 2009). Research on very young children’s impressive linguistic abilities raises questions about how much social information children need before they can correctly interpret new words (with the answer often being: not very much, Hoff & Naigles, 2002). Specific computational tools that children might use to acquire language will be discussed further below.

Gleitman and her colleagues demonstrated the need for linguistic data rather than simply social data in order to compute the meaning of a word by experimentally simulating a child’s word learning environment and presenting it to adults (referred to as the Human Simulation Paradigm; Gillette, Gleitman, Gleitman, & Lederer, 1999). This is an interesting take on language acquisition research, and assumes that adults come to the learning experience with similar abilities to children, if not more or better abilities. In these studies adults watched naturalistic scenes of parents and children interacting, and heard some of the dialog between the people in the scene. One particular word in each scene was replaced with a tone, and the participants had to guess what word was being used. Varying amounts of dialog were silenced to assess the amount of linguistic information necessary to discover a lexical meaning. In order to correctly guess a missing verb from a video of a social situation, adults needed a great deal of linguistic information; that is, they needed much of the dialog to be audible in order to guess the verb (Gillette et al.; Snedeker & Gleitman, 2004; Gleitman, Cassidy, Nappa, Papafragou, &

Trueswell, 2005). While the linguistic information comes from the “outside” in this case, it is being processed by aspects of the person interpreting the word, rather than the word meaning being provided by the social context at hand. If the social situation, including gestures, facial expressions, and actions, were enough to pick up on the lexical meaning of the word, it would be quite easy to guess without much dialog in place. However, the words around the target verb, including their meaning, part of speech, and intonation, were necessary to compute the potential meaning of that target. Under the belief that adults should be able to do as well or better than children (not a far reach, given the social and metalinguistic knowledge that adults gain from experience), these data suggest that children need more than a rich social environment to learn a lexical item.

Taken together, data that support computational theories show children as language learners that exhibit expertise far beyond what is available to them in their social environment, and suggests that children have domain-specific abilities that are used when acquiring language.

An unsatisfying dichotomy

The two schools of thought described above posit different core abilities in children that lead to mastery of a language. The social theories seem insufficient due to their emphasis on context as a cue to lexical meaning in light of the data from the Human Simulation Paradigm studies (Gillette et al., 1999; Snedeker & Gleitman, 2004), and seem insufficient in explaining grammatical development due to the speed and accuracy with which children combine words. Computational theories, however, seem unsatisfactory in that they do not incorporate the demonstrated effects of joint attention on language development, but rather emphasize abstract rule learning as key. Given that neither group of theories gives a satisfactory description of the language development process, it seems clear that language acquisition and development incorporates both social and computational aspects. Thus far, however, the theories and experimental paradigms have not examined how the two are weighted over the course of development. Therefore, this dissertation will compare the effects of certain social and computational abilities on measures of language growth. One important aspect of the social theories is joint attention, which is particularly impaired in children with ASD. In order to compare theories and assess language development across groups of children, joint attention was used as a predictor of language growth. Important aspects of the computational theories are the implicit biases based on linguistic or perceptual information that children seem to use in determining the meanings of new words. Measures of shape bias, word order comprehension, and noun bias, and vocabulary size were used in the following analyses as measures of computational language ability.

Sources of information in language development

Two people are said to be engaged in joint attention when they are both attending to the same object and to one another; it is a triadic interaction that included a shared context between two people and an object on which their attention is focused (Moore & Dunham, 1995). This is an important ability that children typically begin to develop at around nine months of age, and it allows them to expand their world of experiences beyond dyadic interactions between themselves and another person or themselves and an object (Tomasello, 1999). This opens the door to learning about people and objects in new ways, and has been of particular interest to researchers looking at language acquisition and development. There is an intuitive relationship between sharing another person's attention with an object and discovering a word for that object, demonstrated dramatically and compellingly in the play about Helen Keller (Gibson, 1957), whose disabilities in sight, hearing, and speech resulted in an extreme difficulty in establishing joint attention; and demonstrated empirically in a number of studies, some discussed previously (e.g., Akhtar et al., 1991; Carpenter et al., 1998).

Computational theories lean less heavily on the idea of joint attention as a core ability in language acquisition, but leave some room for its role. Hoff and Naigles (2002) directly compared the contributions of shared social context and computational abilities in vocabulary development. They looked at word learning in 18-29 month olds in a naturalistic setting and found that their lexical ability was not as dependent on joint attention (one of several social communication abilities they measured) as on computational aspects of their language input, such as the number of words they heard and the complexity of utterances. Hoff and Naigles allow for the possibility that the very earliest stages of word learning (i.e. in very young children) may depend more on joint attention than later stages.

In an experimental setting, Hollich, Hirsh-Pasek, and Golinkoff (2000) found that younger children were *less* likely than their older counterparts to use joint attention cues to a novel word's meaning. At first glance these findings seem contradictory, but the settings were different: Hoff and Naigles (2002) recorded naturalistic interactions while Hollich et al. used an experimental paradigm. Furthermore, the children in Hollich et al.'s study were younger than the children Hoff and Naigles studied, and younger than those children that Baldwin (1993) found were able to read gaze and intention in a word learning situation. This may indicate that very young children have difficulty using joint attention on their own for word learning, but that parents step in to help children jointly attend and learn vocabulary in naturalistic settings. These studies shed some light on the role of joint attention in word learning relative to computational abilities, but leave open questions about grammar learning, and about children who have trouble with joint attention.

What happens, then, when a child is unable to jointly attend? Are some children who develop language atypically the same as those who have difficulty with joint attention? This is where the experience and behaviors of children with ASD become important. Children with ASD show deficits in joint attention (Mundy, 1995; Tek, 2010), and many show deficits in language as well (Tager-Flusberg, Calkins, Nolin, Baumberger, Anderson, & Chadwick-Dias, 1990; Siller & Sigman, 2008; Eigsti et al., 2007; Boucher, 2012).

The computational abilities assessed in this study include two conceptual biases and one grammatical ability. The conceptual biases are noun bias and shape bias measures. A noun bias refers to the idea that typically-developing children not only have a tendency to acquire nouns first in their lexicon (Gentner, 1982; Gentner & Boroditsky, 2001), but also that they assume a novel word refers to an object rather than an action (Golinkoff, Mervis, & Hirsh-Pasek, 1994; Waxman, 2004). There is evidence that children with ASD use this bias as well in determining word meanings (Swensen, Kelley, Fein, & Naigles, 2007). The shape bias comes into play as children are compiling a vocabulary of nouns; they tend to assign a novel noun to the shape of an object rather than other properties of the object, such as color or texture (Landau, Smith, & Jones, 1988; Smith, 2000). A recent study of the shape bias in children with ASD suggests that these children do not use shape as a clue to word meaning like typically-developing children do (Tek, Jaffery, Fein, & Naigles, 2008). Another ability that typically-developing children use to discover word meaning is word order, or phrasal syntax (Hirsh-Pasek & Golinkoff, 1996). Children pick up on the fact that English sentences follow a subject-verb-object (SVO) pattern, and can identify the meaning of a new word depending on its place in a sentence with relation to other words. Syntactic bootstrapping is another way that children use syntax to learn word meanings; they can narrow down the potential meanings of a novel word based on the other parts of the sentence. The order of noun phrases around a verb is instructive as to whether that verb is causal or not. For instance, “The duck is *kradding* the bunny” indicates a causal action while “The duck and the bunny are *kradding*” does not. This ability has been demonstrated in typically-developing children as young as two years of age (Naigles, 1990). There is some evidence that children with ASD use these tools like TD children (Kjelgaard & Tager-Flusberg, 2001; Shulman & Guberman, 2007; Naigles, Kelty, Jaffery, & Fein, 2011). Finally, an estimate of the simple number of words in a child’s vocabulary was used as a computational measure, with the idea that early vocabulary size is a source of data for the child to use when developing later language skills. For instance, using SVO word order or syntactic bootstrapping depends on the child’s prior knowledge of other words in the sentence.

The current study analyzed the ways that social and computational variables affect growth of some specific language measures. The measures chosen for these analyses were children’s mean length of utterance, the number of words children used during free play, the number of pronouns children used during free play, and the progression of their shape bias ability. The latter is a computational ability described above, and its progression over time was used as a dependent variable in these analyses. All of these dependent measures will be discussed in more detail at the end of the introduction. These measures were tracked over time in order to identify important predictors in their development. Mean length of utterance is an early measure of grammatical complexity (Brown, 1973), and number of word types is an early measure of lexical ability. Pronoun use is a measure of a particular lexical skill, as children need to learn how to use particular words in changing ways depending on their relationship to a communication partner (i.e., *you* versus *I*). Examining the development of shape bias with relation to social and computational abilities could illuminate the contributors to this early-used lexical bias. Furthermore, using a longitudinal method to analyze these variables allowed us to ask questions about how the pattern of growth over time might predict later language and cognitive skills. As discussed earlier, the complicated abilities and disabilities of children with ASD made them an important population to compare with typically-developing children.

Autism spectrum disorders and language

ASDs are a heterogeneous range of neurodevelopmental disorders that are characterized by deficits in language and social abilities and by restricted, repetitive behaviors (APA, 2000). Part of the diagnosis of classic autism involves a delay in the onset of speech, and older children with autism show deficits in many areas of language, from morphosyntax to pragmatics (Eigsti et al., 2007). While 25% of children with autism may never gain functional speech (Lord, Risi, Lambrecht, Cook, Leventhal, DiLavore, Pickles, & Rutter, 2000), many do, and some may reach close to normal language abilities (Boucher, 2012). Children with autism are identified by atypical socialization and communication, and stereotyped or repetitive behaviors (Levy, Mandell, & Schultz, 2009). They also show late onset of spoken language, with some never producing spoken language (Gerenser, 2009).

How might these atypicalities affect language development? Children with autism show significant deficits in joint attention development. Children with ASD are, by diagnosis, socially impaired. In particular, young children have difficulty following eye gaze, pointing and following points, and imitating other's behaviors; the types of behaviors that typically-developing children are engaging in beginning at around nine months of age (Mundy, 1995; Gerenser, 2009). There are studies suggesting a specific link between these social abilities and the rate of language growth in children with ASD. Siller and Sigman (2008) measured the joint attention of 28 young children with ASD using experimental procedures specifically designed to elicit joint attention. They also assessed language abilities at four times over about four years, and found that children's response to joint attention, as well as and various measures of maternal synchronization with her child's actions and language, were strong predictors of the rate of language gain. Charman, Baron-Cohen, Swettenham, Baird, Drew, and Cox (2007), Paul, Chawarska, Cicchetti, and Volkmar (2008), and Anderson, Lord, Risi, DiLavore, Shulman, Thurm, Welch, and Pickles (2007) have all demonstrated that joint attention abilities are significant predictors of later language outcome in children with autism. The link between joint attention deficits and language was also demonstrated in a study showing that infants and school-aged children, both with autism, showed deficits in attending to a speaker's direction of gaze when it was providing clues as to the meaning of a novel word. Children with autism were less able to correctly assign the word to the object than children with an unspecified mental handicap who were matched on age and other cognitive measures (Baron-Cohen, Baldwin, & Crowson, 1997).

In spite of these joint attention deficits, the majority of children with ASD do learn language. Children with ASD have been thought to have major deficits in the pragmatics of language, but little difficulty with syntax (Tager-Flusberg, 1994). Indeed, children with ASD have demonstrated use of syntax to interpret new verbs in a syntactic bootstrapping paradigm (Shulman & Guberman, 2007; Naigles, et al., 2011). More recently, though, evidence has shown some grammatical impairment in autism, including generally less complex utterances as well as inconsistent morphological abilities (compared to developmentally delayed and mental-age matched TD children; Eigsti et al., 2007). Park, Yelland, Taffe, and Gray (2012) studied morphology use in children with ASD and found mixed results, with the ASD group performing similarly to the TD group on some grammatical rules (e.g. articles, negations), but more like a developmentally delayed group on other aspects (e.g. past tense endings). Roberts, Rice, and Tager-Flusberg (2004) also found impairment of tense marking in some subgroups of children with ASD. Overall, the results are mixed and call for a more detailed longitudinal approach to examine not only the outcome of lexical and grammatical development in ASD, but also the rate and trajectory as compared to typical development.

Tek, Mesite, Fein, and Naigles (2013) have completed such a study with some of the same children who are included in this dissertation. Longitudinal analyses of spontaneous speech supported the idea that some children with ASD are similar to those with typical development, while others are globally impaired. The HVA children in their study showed some slower growth in specific language areas like MLU, plural markers, and the third-person irregular verb tense, but were dramatically more similar to the TD group than the children with LVA. These are the types of findings that can come out of a longitudinal study of specific language measures, and this dissertation will expand on such findings to include predictors of interest that were not included in Tek et al.'s study.

With regards to computational abilities, children with ASD are wide-ranging. A small minority of this group show outstanding abilities in some particular skill, in many cases involving math or computation (Treffert, 2009), while others show IQ's consistent with mental retardation (Fombonne, 2005). The computational word-learning biases described earlier (noun bias, shape bias, and word order) are implicit biases that involve abstracting patterns out of specific examples; for instance, English-acquiring children hear SVO sentences over and over in their input and come to understand that novel strings of words can be interpreted as SVO. This type of abstraction is one that people with ASD are thought to have particular difficulty with (Happé & Frith, 2006). However, Eigsti and Mayo (2011) reviewed evidence from numerous studies of implicit learning in adults and children with ASD, and found mixed results. In some cases, those with ASD performed equally to typical individuals, and in others, relatively subtle deficits were demonstrated. Specifically, school-aged children with ASD showed subtle deficits in an artificial grammar-learning paradigm similar to the one used by Saffran et al. (1996); this could be a crucial ability in learning grammatical patterns (Mayo & Eigsti, 2010). The wide variety of skill and ability level in computational abilities leaves open the question of the contribution of these abilities to language development.

According to social theories of language development, children with deficits in joint attention should suffer overall deficits in language acquisition and development. A strict computational theory might suggest that children with impairments in joint attention but intact analytical abilities would show unimpaired language progress. In children with ASD, we have a group who are impaired in joint attention and variable in computational abilities, and show a complex pattern of language abilities. Muddling the questions and findings in this field of research are distinctions made between lexical development and grammatical development. Joint attention is fairly well-established as a tool for lexical learning (e.g., Baldwin et al., 1996), but less well-established is the relationship between social abilities and grammatical learning. Even in the typically-developing literature few direct relationships between early social ability and later syntactic ability are found. Rollins and Snow (1998) did a longitudinal study of six children with autism and found that individual growth in grammar was predicted by the number of shared attention episodes between child and parent. Park (2013; Park, Tek, Fein, & Naigles, 2012) found positive predictive relationships between early joint attention in children with ASD and later syntactic bootstrapping and wh-question performance, both of which require some grammatical skill. These studies suggest some positive relationship between social ability and grammatical ability in children with ASD, but further study is warranted, and using a longitudinal perspective could help identify specific ways that social abilities serve later grammatical skill.

A developmental approach

There is a long history in psychology and linguistics of trying to capture children's language development longitudinally in order to assess the pattern of acquisition and processes of development (e.g., Brown, 1973; Tomasello, 1992b; Naigles, Hoff, & Vear, 2009). Adolph, Robinson, Young, and Gill-Alvarez (2008) made a convincing argument for the case of sampling children's abilities frequently over time in order to identify developmental trajectories, rather than simply looking at early and late abilities. Without frequent sampling, development can appear step-wise and information about the patterns can be lost[1]. In cases where development is atypical, the analysis of trajectories and developmental patterns can be crucial for understanding the disordered development (Karmiloff-Smith, 1998; Thomas & Karmiloff-Smith, 2003). Disorders like ASD are developmental in nature, meaning that they may manifest differently across time, and the key to identifying important differences may lie in the trajectories of development rather than in the outcomes. Karmiloff-Smith suggests that cognitive disorders involve a continuous spectrum of abilities, and that consideration of the process of development is necessary to understand a disorder. As mentioned earlier, research into language development in ASD has revealed a complicated picture of language acquisition, and following Karmiloff-Smith's line of thinking, it could be that very subtle differences in children's experience lead them down very different paths of development. Are certain children following a similar

[1] Adolph et al. (2008) studied motor abilities, and their conclusions about ideal sampling windows are smaller than the windows used in this study. Despite these differences, the basic idea holds: that a frequently-sampled longitudinal design gives a better view of developmental trajectories than a cross-sectional or two time point design.

trajectory as TD children? Are children with ASD simply slower than TD children, or are they following a dramatically different path? How does change over time relate to later skills and abilities that the child might exhibit? A developmental approach allowed us to research these questions, rather than simply analyzing cross-sectional outcome data from children with typical and atypical language development.

Some previous research into language development in ASD has taken a rigorous longitudinal approach for similar questions. Siller and Sigman (2008) analyzed language abilities at three timepoints and found that joint attention was important to the trajectory (not just outcome) of language growth. Tek and colleagues' (2013) use of longitudinal analyses revealed two subgroups within a group of children with ASD with regards to morphological development; one group was more similar to TD controls while another group had significantly flatter trajectories. Goodwin, Fein, and Naigles (2012) saw that children with ASD followed a similar process of wh-question acquisition to TD children although their development was delayed. The comprehension of those questions by children with ASD preceded their production; a conclusion which could only be reliably reached by looking at tightly-spaced longitudinal data. In language development research with typically-developing children, Rowe, Raudenbush, and Goldin-Meadow (2012) used longitudinal methods to create models of vocabulary growth based on children's characteristics (socioeconomic status and gesture use) and found that the change over time during an early time period predicted later vocabulary ability. That is, the model of change in vocabulary when the children were toddlers was a better indicator of school-age vocabulary than a simple one-time early measure of vocabulary. These longitudinal designs provided us with important groundwork for the current study, which uses a longitudinal design and data analysis methods.

The current study

To examine the patterns of and contributors to growth in language measures of children with ASD and those who are typically developing, we followed two groups of children over the course of four years, and collected naturalistic and experimental data on their social and computational abilities. We used longitudinal modeling techniques to examine how different variables contribute to children's language growth.

We chose four measures of interest to model for this project. They were chosen to cover a range of levels of specificity in language use, and to sample from both language comprehension and production. Two of the measures, mean length of utterance and number of word types, offer us general measures of language production from grammatical and lexical perspectives. Brown (1973) introduced the idea of measuring a child's grammatical complexity by averaging the number of morphemes the child uses per utterance. Brown found that MLU was a better indicator of grammatical ability than something as simple as a child's age, because while children develop grammatical abilities at slightly different ages, they tend to go through a phase of producing one morpheme utterances, followed by a rapid increase of combinations. MLU has been used extensively as a measure of grammatical complexity, and in this study we used it across time to track grammatical development.

We also wanted to track children's lexical, or vocabulary, development. Some frequently used measures of vocabulary are based on parent report (e.g. CDI; Fenson, Dale, Reznick, Bates, Thal, & Pethick, 1994), but we chose to use a speech sample to extract our measure of vocabulary growth. We tallied the number of *different* words a child used during our speech sample. This is referred to as the number of word types a child produced. The number of tokens is the total number of words that a child produced during that time, which we did control for in our analyses, as discussed in the method and results section.

Pronoun use is a particular area of interest in the language development of children with autism. Personal pronouns (particularly *I*, *me*, and *you*) are conceptually interesting because they require speakers to adjust their use based on their relationship to the subject. Children must understand that when their parent addresses them as “you”, they must respond using “me” if they are to refer to the same person correctly. English-speaking children begin producing personal pronouns (*I*, *me*, *you*, *he*, *she*) around the age of 24 months (Cruttenden, 1977), and use them fairly accurately. There are some cases in which children reverse pronouns and use “I” for “you” and vice versa, but current findings on the prevalence of pronoun reversals in children with ASD are mixed (Naigles, Cheng, Khetrapal, Rattanasone, Fein, & Demuth, submitted). One important finding in the literature about pronoun use in ASD is that children tend to use more nouns and proper names than pronouns when compared with TD children (Jordan, 1989; Lee, Hobson, & Chiat, (1994). This is despite the fact that pronouns are often shorter than nouns and proper names, and thus may be easier to produce. It is likely that there are conceptual and social challenges to using pronouns correctly, due to their changing referents depending on the perspective of the speaker (Loveland & Landry, 1986; Evans & Demuth, 2012). In addition, general language skills (rather than specific social abilities) may affect the pronoun acquisition of children with ASD (Naigles et al., submitted). We wanted to tap into a very specific area of language development in children with ASD, and so we tracked the number of personal pronouns that children used across time (whether used correctly or incorrectly).

These first three measures of interest (MLU, number of word types, and number of personal pronouns) were all taken from speech samples. Previous work using speech samples have used smaller groups of children and shorter periods of speech than the current study (Rollins & Snow, 1998; Tager-Flusberg, 1994). Many other studies (e.g., Siller & Sigman, 2008) are based on standardized measures of language, which depend heavily on elicited tasks or parent report. By using spontaneous speech samples over time we are able to capture a moment in the child's naturalistic language development without depending on elicited speech, which may be difficult for children with ASD due to their difficulty with social interactions (APA, 2000).

The last measure we included was a measure of a particular lexical bias. The shape bias has been studied extensively as a way that children (and adults) narrow down potential meanings of a new word (Landau et al., 1988; Smith, 2000; Abecassis, Sera, Yonas, & Schwade, 2001; Tek et al., 2008). Adults and children (after building up a vocabulary of nouns) tend to assume, when hearing a new word in the presence of a new object, that the word refers to that object's shape rather than color, texture, or size. This is a powerful tool in learning new words, because it narrows the potential options to consider when trying to interpret a new word. Assessing the shape bias gives us a window into one of the *processes* of language acquisition, and in particular word learning, not just the *outcome*. The process of language acquisition in children with ASD is of much interest from both a clinical perspective and a research perspective. For this reason, we tested children's use of the shape bias over time to tap into their underlying processes of word learning. This measure was tested using looking time (described in more detail in the Method section) which reduces the demands on young children (no matter their developmental ability) and gives us a window into implicit computational processes.

Our four measures of language development over time, then, span production (MLU, word types, pronouns) and comprehension (shape bias) at the lexical (word types, pronouns, shape bias) and grammatical (MLU) levels. The diversity of measures is an important step forward in the longitudinal study of language acquisition in both a typical and atypical population.

Prospectus

Four measures of early language abilities were tracked over time in two groups of young children: one with ASD and one TD. Social and computational abilities in these children were assessed at several timepoints and their contribution to the growth in the four language measures was assessed. Additionally, the patterns of growth in these four measures were tested to see if they predicted later language and cognitive outcomes. The contributions of social and computational abilities were assessed with regards to the developmental trajectories of the four measures.

2. Method

Participants

All children in this study were visited every four months in their homes over two years, making up a series of six visits (V1-V6). Two years after V6, a follow-up visit (V7) was completed. A total of 18 typically-developing children and 22 children with ASD participated in the study, with slightly different subsets included for different analyses; the specifics are outlined in the next section. Children were recruited through word of mouth, mailing lists, and autism service providers. The group was representative of the populations in Connecticut, Massachusetts, New York, Rhode Island, and New Jersey from which they were recruited, and the gender imbalance is based on the higher prevalence of ASD in boys than girls.

The children with ASD had been diagnosed by a clinician within six months of the first visit, and diagnosis was confirmed through administration of the Autism Diagnostic Observation

Schedule (ADOS; Lord et al., 2000) and the Childhood Autism Rating Scale (CARS; Schopler, Reichler, & Renner, 1988). The TD children were younger than the children with ASD in order to match their overall language abilities at V1; there was no difference between the groups' mean lengths of utterance or their vocabulary size as measured by the CDI. There were, of course, significant group differences in their ADOS, CARS, and Vineland Socialization scores due to the characteristics of children with ASD. Matching our samples on language ability at V1 allowed us to examine language development starting from a similar point regardless of overall initial delay. Table 1a displays demographic and diagnostic information for all participants.

For some of the analyses, the children with ASD were split into two groups, in order to address the considerable debate over the degree to which individual children with ASD are deviant in their language development. We split the ASD group based on their V1 score on the Mullen Expressive Language test (Mullen, 1994). A median split of the children for whom we have the full battery of tasks gives us 11 children we will refer to as children with high verbal autism (HVA), and 11 with low verbal autism (LVA) (Tek et al., 2013). We refer to this way of sorting the children as subgroup, with three levels (TD, HVA, and LVA), while diagnosis refers to leaving the children with ASD as one group. Table 1b displays the characteristics and group differences for the subgroups.

Table 1a. Demographic and diagnostic information at V1. ASD and TD columns contain means and standard deviations; t-value and p-value columns demonstrate whether there are group differences.

	ASD	TD	t-value	p-value
N	22 (1 female)	18 (2 female)		
Age in months	33.09 (4.26)	20.68 (1.88)	-11.47	<.001**
ADOS	13.82 (3.86)	0.11 (.32)	-14.98	<.001**
CARS	34.77 (6.35)	15.39 (.76)	-12.85	<.001**
Vineland Social	74.91 (7.58)	100.5 (7.0)	10.99	<.001**
CDI words	113.95 (118.95)	118.78 (114.35)	0.13	.89
Mean Length of Utterance	1.5 (.68)	1.4 (.25)	-0.73	.47

Table 1b. Diagnostic information for the subgroups within the ASD group.

	HVA	LVA	t-value	p-value
N	11 (0 female)	11 (1 female)		
Age in months	31.14 (3.35)	35.03 (4.30)	-2.37	.03*
ADOS	11.73 (2.87)	15.91 (3.67)	-2.98	.007*
CARS	30.36 (5.43)	39.18 (3.52)	-4.52	<.001**
Vineland Social	77.09 (8.36)	72.73 (6.36)	1.38	.18
CDI words	200.27 (105.42)	27.64 (46.96)	4.96	<.001**
Mean Length of Utterance	1.95 (.75)	1.10 (.13)	3.71	.001**

Procedure

Each child was visited at his or her home by two to four experimenters. The experimenters administered a series of standardized tests, some of which involved eliciting and observing the child's behaviors and others that involved questions for the parent. The particular battery of tests differed slightly between visits depending on age-appropriateness and time constraints. After the tests, children watched a series of videos that assessed language comprehension using Intermodal Preferential Looking (IPL; Hirsh-Pasek & Golinkoff, 1996), which is a simple and non-social task well-suited to children with both typical and autistic development. IPL uses a series of side-by-side video clips to see whether children are following audio cues that instruct them to look at particular scenes. In this study, a portable screen and laptop with projector were set up to show side-by-side video clips, and a video camera recorded the child's face as he or she watched the screen. In the lab, a custom program was used to code the video of the child's face frame-by-frame for their direction of gaze. Comprehension was assessed by analyzing the amount of time a child looked at each side of the screen (Naigles & Tovar, 2012). Specific examples of how IPL assesses language comprehension are described below.

After watching the videos, the child and his or her parent (the mother in all cases except two) were videotaped while playing together for 30 minutes. A bag of toys – the same for all participants at each visit – was provided, and for 15 minutes the parents was provided with prompts to engage the child in certain activities (Stone, Coonrod, & Ousley, 2000). The remaining 15 minutes were spent in free play. In the lab, all of the speech was transcribed and analyzed using CLAN (MacWhinney, 2000). The analysis revealed counts of words and different parts of speech. All of the measures were extracted from these activities as described below.

Measures

There were three different groups of measures used in these analyses. One group was made up of four *measures of language growth*: mean length of utterance, number of word types produced, number of pronouns produced, and comprehension of the shape bias. These measures were taken at multiple visits, and thus we can model their change over time. Another group was made up of *predictors taken at V1*. These measures were initial assessments of language and cognitive abilities that were used to explain variance in the change over time of the four measures of language growth. The third was a group of *measures taken at V7*, which will be referred to as outcome measures. The V7 measures were included to see if they were predicted by the models of the four measures of language growth.

Measures of language growth

Mean length of utterance (MLU): This measure was calculated for the 18 TD children, 10 of the HVA children, and 7 of the LVA children. The transcribed speech from the play session was analyzed for the number of morphemes used in each utterance. For instance, “Hi mama” contains two morphemes, while “I like the elephants” contains five morphemes (*elephant* being one morpheme and the plural -*s* being a second). The total number of morphemes a child produced over a 30-minute session was divided by their total number of utterances over that session. This provided a measure of grammatical complexity (Brown, 1973).

Word type count: This measure used the same groups as the MLU analysis. The number of *different* words a child produced during a 30-minute play session was calculated using CLAN. This is distinguished from the number of words overall, referred to as tokens. For instance, saying the word “dog” five times during a session results in one word type, but five word tokens. Word types and word tokens are measures of the diversity of words a child knows and their talkativeness, respectively. They have been shown to be highly correlated (Hoff & Naigles, 2002), but were used separately in these analyses; their roles in the analyses will be discussed below.

Pronoun count: This measure used play session data from 18 TD children and 15 children with ASD; the ASD group only included the children with ASD who produced any pronouns at any visit. This criterion necessarily excluded most (but not all) of the LVA group. We analyzed the total number of pronouns that a child produced during each 30-minute session, including first-, second-, and third-person pronouns, but only those that were deemed to have an unambiguous referent by a trained coder (Naigles et al., submitted; Cheng, Khetrapal, Demuth, Fein, & Naigles, 2012). This also included pronouns that were used incorrectly by the child, which made up a very small proportion of the overall number of pronouns.

Shape bias comprehension (Tek et al., 2008): This measure was calculated from 16 TD children and 14 children with ASD, based on data that were available at the time of analysis. The shape bias is a tendency to guess that a novel word refers to the shape of an object rather than other properties (Landau et al., 1988). The shape bias comprehension measure was taken from an IPL video that children watched during V1 through V4. Five target objects and 10 test objects were created using blocks and Legos. Each of the target objects had one same-color test object and one same-shape object. The children saw two blocks of videos: a no-name block and a name block. The no-name block showed each test object accompanied by the audio "Look at this!" and then showed the two test objects side by side with the audio "Which one looks the same?" The name block showed each target object accompanied by an audio such as "Here's the dax!" and then showed the two test objects side by side and asked "Where's the dax?" See Table 2 for a full layout of the video.

We calculated the difference between the time children spent looking at a target object during the ‘name’ trial and the ‘no name’ trial. This difference score captured the amount of time they spent looking at the same-shape item during the name trials compared with the no-name trials. Positive values meant that the child looked longer at the same-shape item when it was named, demonstrating use of a shape bias to extend a novel noun, and negative values indicated that the child looked longer at the same-color item during the name trial than the no-name trial, which we took to be a lack of use of the shape bias.

Table 2. Layout of video used to test shape bias via comprehension

	Video 1	Audio	Video 2
No name trials			
1	Wooden u-shaped block	Look at this!	Blank
2	Blank	Look at this!	Wooden u-shaped block
3	Wooden x-shaped block	They’re different now!	Sparkly u-shaped block
4	Wooden x-shaped block	Which one looks the same?	Sparkly u-shaped block
Name trials			
5	Wooden u-shaped block	Here’s the dax!	Blank
6	Blank	Look, a dax!	Wooden u-shaped block
7	Wooden x-shaped block	They’re different now!	Sparkly u-shaped block
8	Wooden x-shaped block	Where’s the dax?	Sparkly u-shaped block

Predictor measures at V1

ADOS: The Autism Diagnostic Observation Schedule (Lord et al., 2000) is a diagnostic tool for ASD. A trained research assistant engaged the children in a series of activities that were designed to identify deficits in social and communicative behaviors. A score of higher than 7 indicates a disorder on the autism spectrum, and higher than 12 indicates classic autism. At V1 it was used simply to verify the diagnosis already given by a clinician. Apart from these two cutoff scores, the score on the ADOS does not indicate severity of autistic symptoms; it is not considered to be a continuous variable.

Mullen: The Mullen Scales of Early Learning (Mullen, 1994) provided scores for language and motor abilities through a series of activities that were administered by a trained research assistant. It is commonly used for assessing both typical and atypical development. The expressive language subscale was used for dividing the children with ASD into HVA and LVA groups. The subscale includes, but is not limited to, assessment of babbling and vocalizations, naming pictures of household objects, and repeating sentences. We also used the Visual Reception scale as a measure of non-verbal intelligence. The Visual Reception scale includes attention to pictures, matching and sorting objects, and simple memory tests.

Vineland: The Vineland Adaptive Behavior Scales (Sparrow, Balla, & Cicchetti, 1984) is a semi-structured interview conducted with a caregiver (in these cases usually the mother) that scores children’s communication skills, daily living activities, socialization abilities, and motor skills. The socialization subscale was used in these analyses; the parent was asked questions about whether the child showed preferences and affection for certain people, demonstrated a desire to please others, imitated complex actions, or used emotion words.

CDI: The MacArthur Communicative Development Inventory (Fenson et al., 1994) is an assessment of vocabulary size. Parents were provided with a list of common words that occur in children's early vocabulary, and they checked off the words that their child produced.

Word order percent looking and word order latency: These measures were based on another IPL video that children watched at V1. Two characters in these videos performed reversible actions (e.g. tickling, pushing) and children were instructed to look at one of the scenes. The difference between the two scenes depended solely on word order (e.g. 'girl tickling boy' vs. 'boy tickling girl'). Six verbs (*tickle, ride, kiss, hug, push,* and *wash*) were tested in this way. A sample layout is presented in Table 3. One measure analyzed from this task was the difference between the percent of time that a child spent looking at the correct scene during the directing audio (step 3 in Table 3) and the percent of time he or she spent looking at the correct scene during the non-directing audio (step 4 in Table 3). This was referred to as 'word order percent looking'. Another measure was simply the amount of time between the onset of the video in step 4 of Table 3 and when the child looked to the correct scene. This is referred to as word order latency (Naigles et al., 2011).

Noun bias percent looking and noun bias latency: This IPL video tested children's tendency to extend a novel word to an object rather than an action at V1. An unusual puppet performed an unusual action, and children heard an ambiguous novel word that accompanied the scene. The novel word ended in *-en*, so that it could be interpreted as a noun (like *kitten*) or a verb (like *walkin'*). The puppet was then shown performing a different novel action (same-object scene), and a different puppet was shown performing the same action (same-action scene). Children were asked to find the novel verb, meaning that they had to look at the scene that contained the aspect they assigned to the novel verb. See Table 3 for full layout. Six different novel words were tested in this way. The noun bias percent looking measure was calculated by finding the difference between the percent of time a child spent looking at the same-object scene during the non-directing audio and the directing audio (steps 4 and 5 in Table 3). Positive numbers indicated that a child spent more time looking at the same-object scene during the directing audio, and we interpreted that to mean the child used a noun bias. The noun bias latency measure is the amount of time between the beginning of step 5 and when the child looked at the same-object scene (Swensen et al., 2007; Tek et al., 2008; Naigles et al., 2011).

Joint attention measures: Joint attention was assessed based on the play sessions (Tek, 2010). Trained coders watched the videos of the play sessions and marked the beginning and ends of joint attention sessions, and who initiated those sessions. Response to joint attention (RJA) involved the duration of time a child spent in joint attention when the child had responded to a bid from the parent to attend to a toy. The duration of RJA at the play session during V1 was used as a predictor in some of the following analyses, and is referred to as 'initial RJA'. Another predictor is called 'RJA change'. This value was calculated by finding the average slope of RJA duration during V1 through V4, and then finding the difference between that average slope and each individual child's slope. So, each child was assigned a particular slope for their change in RJA from V1 to V4, and their 'RJA change value was the difference between his or her individual slope and the average slope. If a child had a negative value on this measure, it meant that his or her slope was shallower than the average slope, and a positive number meant that the individual's slope was steeper than the average slope. Initiation of joint attention (IJA) was recorded when the child made a bid for attention and the caregiver followed that bid, but there were such low levels of IJA in any of the children that there were no analyses performed using IJA.

Table 3. Layout of videos used for word order and noun bias comprehension measures

	Video 1	Audio	Video 2
Word Order			
	Girl waves	Look!	Blank
	Blank	Look!	Boy waves
	Girl waves	Look!	Boy waves
	Girl waves	Where is the girl?	Boy waves
	Girl waves	Where is the boy?	Boy waves
1	Girl tickles boy	Look, tickling!	Blank
2	Blank	See, tickling!	Boy tickles girl
3	Girl tickles boy	Hey, tickling!	Boy tickles girl
4	Girl tickles boy	Look, the girl is tickling the boy!	Boy tickles girl
Noun bias			
1	Possum puppet digs with nose	Here's *toopen*!	Blank
2	Blank	See, *toopen*!	Possum puppet digs with nose
3	Possum puppet digs with nose	Look, *toopen*!	Possum puppet digs with nose
4	Possum puppet sways side to side	They are different now!	Beetle puppet digs with nose
5	Possum puppet sways side to side	Where's *toopen*?	Beetle puppet digs with nose

These predictors were grouped into computational measures and social measures of functioning as shown in Table 4. As discussed in the introduction, a major issue addressed in this study is the role of social and computational abilities in language growth. The ADOS and Mullen Visual Reception scores are not included in the table, because they were used only as a diagnostic measure or as a grouping measure, rather than as a predictor of future growth.

Table 4. Variables organized by type of measure. All measures are for V1 unless otherwise indicated.

Type	Variable Name	Description
Computational	CDI	Number of words known on the CDI
	Word Order % Match	Difference in percent of time spent looking at match minus non-match for word order video
	Word Order Latency	Amount of time before first look to match for word order video
	Noun Bias % Match	Difference in percent of time spent looking at match vs. non-match for noun bias video
	Noun Bias Latency	Amount of time before first look to match for noun bias video
Social	Initial RJA	Time spent in joint attention episodes which begin by responding to parental bid for joint attention
	RJA change	Difference between individual’s slope and average group slope in RJA increase over first four visits
	Vineland Social	Score on Vineland socialization test

Outcome measures at V7

The Vineland and the ADOS were re-administered at V7, along with the following tests. These tasks are outlined in Table 5.

TACL: The Test for Auditory Comprehension of Language (Carrow-Woolfolk, 1985) is a test of receptive language abilities divided into vocabulary, morpheme, and elaborated sentences subtests. The test gave us scores for a child's vocabulary, morphology, and syntactic abilities.

DAS: The Differential Ability Scales (Elliot, 2007) is an IQ test for children that includes working memory, picture and shape matching, and delayed recall.

Wug task: The wug task (Berko, 1958) is a test of productive morphology. Children heard 20 sentences and had to fill in the "blank," such as "This is a wug, this is another wug, now there are two…" with the correct answer being "wugs". This task tested plurals, past and progressive tenses, third person singular conjugations, and possessives using novel words. Each sentence was accompanied by a picture with ambiguous figures and settings to assist the children in producing a word for the "blank".

Mental verb comprehension: Children responded to puppets that used the words "think," "guess," and "know" to describe the location of a hidden item (Howard, Mayeux, & Naigles, 2008). For example, the experimenter told the child that one puppet "knew" a sticker was in a blue box, while the other puppet "thought" the sticker was in the yellow box. A steady tone of speech was used throughout to avoid the child using prosodic clues. The word "know" indicated the correct response when paired with either "think" or "guess". "Think" and "guess" were not compared against one another.

Theory of mind: Children engaged in two classic theory of mind tasks. The unexpected location task (Wimmer & Perner, 1983) involved a scene acted out by the experimenter with two dolls and two hiding spots. The dolls placed an item in a hiding spot, and then while one doll was "away and not looking" the second doll changed the hiding spot. In order to answer correctly (and indicate presence of a theory of mind) a child had to tell the experimenter that the first doll still thought the item was in the first hiding spot, while knowing that it actually had changed. For the unexpected contents tasks (Perner, Leekam, & Wimmer, 1987), an experimenter showed a child a crayon box that was actually full of deflated balloons. In order to demonstrate a theory of mind, the child had to answer correctly about what he or she *thought* was in the box before opening it, and what someone else might think was in the box.

Table 5. Visit 7 measures and descriptions.

Variable name	Description
TACL	Standardized language test, including vocabulary, morphology, and syntax scores
DAS	Child IQ test
Wug task	Morphological production task
Mental verb comprehension	Comprehension of "think" vs. "know" and "guess" vs. "know"
Theory of Mind	Unexpected locations and unexpected contents tasks

<u>Analysis method: Growth curve modeling</u>

In order to answer the key questions in this dissertation, we used inferential models of early predictors, which incorporated time as a variable. Growth curve analysis (GCA) is a type of regression that suits this dataset for several reasons. Ordinary least squares (OLS) regression (the type most frequently used) assumes that there is no systematic relationship between variables. However, because this dataset is longitudinal and collected repeated measures from the same participants over multiple times, OLS regression is inappropriate. GCA, on the other hand, treats a time variable ("visit" in this dataset) in a sequential manner, allowing for analyses of *change over time* (Singer & Willett, 2003). Another assumption of OLS regression is homoscedasticity of residuals. When residuals are homoscedastic, the variance has the same shape at all points in the dataset. However, in this dataset the children were young enough at the beginning that they were somewhat similar to one another in their language abilities (i.e., few of them are talking, so the group manifests a small amount of variance). By the later visits, the TD children increased in language abilities, while the ASD group was more varied, with some children improving greatly and others hardly at all. GCA allows for such *heteroscedasticity* without violating any assumptions, unlike OLS. Lastly, *missing data points* are a fact of longitudinal datasets, particularly those with clinical populations. While retention for this study was excellent, there were cases when a child refused to participate in a specific task or when a family was unreachable. GCA is robust enough that data points which are missing at random (i.e., the fact that a data point is missing is not systematically related to a predictor variable) do not adversely affect modeling (Singer & Willett). In this study, GCA was achieved using the software program R (R Core Team, 2013; RStudio, 2012; Bates, Maechler, & Bolker, 2012).

GCA is similar to OLS regression in that it involves creating an equation that includes an intercept added to some series of predictors multiplied by coefficients. A key difference between these two techniques is that GCA includes two levels of equations (Singer & Willet, 2003; Rowe, et al., 2012). See Figure 1, top panel, for an example of what an OLS model for this study might look like including visit and diagnosis group as predictors for a given measure. In this case, Y is the outcome for a particular individual (i), and the outcome for that individual is found by adding an initial intercept (b_0) to relevant variables multiplied by coefficients. A residual error term for the individual (ε_i) is added to make up the difference between the value estimated by the coefficients and that individual's true outcome value.

A growth curve model uses many of the same elements except that there are two levels of equations; one for within-subject variation, and one for between-subject variation. A level-1 model, for this study, includes visit as the within-subjects variable. In each model, we subtracted '1' from visit to create a centered variable so that the intercept actually referred to the calculated value for an outcome measure at visit 1, rather than at some estimated previous time point (i.e. a "visit 0," which did not exist). For that reason, (Visit-1) is used in the equations in Figure 1. The bottom panel of Figure 1 shows the level-1 model, which includes an intercept and coefficient for the visit variable and an error term, and the level-2 model, which includes the diagnosis (between-subject) variable. The different levels of a growth curve model allow for what is called a "random effect" on intercept and slope. The error terms in the level-2 equations (symbol zeta) mean that there is individual variation for each participant feeding into both the intercept and slope variable. Conceptually this differs greatly from OLS regression. While OLS finds an average best-fit line, GCA fits a line for each individual participant on its way to creating the best-fit model[2]. Due to the multi-level aspect of GCA, we can say that we are finding the best-fit growth parameters for our outcome measures, while taking into account individual variation in intercept and slope.

[2] With this type of process there is a chance that a given dataset will be unable to converge on a model. This can happen when there are too few data points, or if the proposed model is too complex for the given data (Singer & Willett, 2003).

Figure 1. Comparison of ordinary least squares and growth curve analysis regression equations.

OLS regression equation:

$$Y_i = b_0 + b_1(\text{Visit}) + b_2(\text{Diagnosis}_j) + \varepsilon_i$$

GCA regression equations:

Level-1: $$Y_{ij} = \pi_{0i} + \pi_{1i}(\text{Visit}_{ij}\text{-}1) + \varepsilon_{ij}$$

Level-2: $$\pi_{0i} = \gamma_{00} + \gamma_{01}(\text{Diagnosis}_j) + \zeta_{0i}$$ $$\pi_{1j} = \gamma_{10} + \gamma_{11}(\text{Diagnosis}_j) + \zeta_{0i}$$

Random effect

Random effect

adapted from Singer & Willett, 2003 and Rowe et al., 2012

The two goals in developing a growth curve model are to find a well-fit model and test hypotheses about predictors and outcomes; these goals must be balanced by the experimenter. Model fit can be mathematically tested (which is explained below), and increasing a model's fit by adding predictors is important. However, solely focusing on model fit can lead to overly complex models that only explain one particular sample, rather than explaining a population (Babyak, 2004). To avoid this, an experimenter can also focus on whether particular predictors of theoretical interest are explaining significant variance, which is not a requisite of increased model fit (i.e., the *combination* of predictors may lead to a good model fit even while *individual* predictors are not significant).

The steps in developing a growth curve model are as follows. First, an *unconditional means model* is used to predict a chosen outcome variable. This model is used to tell whether the outcome variable's intercept is different from zero. For example, a non-zero intercept of a given language measure means that the particular language ability is already exhibited at V1. A second model, called an *unconditional growth model*, includes the visit variable to predict the same outcome variable, and shows whether there is an overall effect of visit (i.e., time) on the outcome. Further steps involve adding main effects and interactions based on hypotheses about the outcome variable. For these models, next steps might include adding diagnosis in interaction with visit, to see if there is an effect of diagnosis group on change over time, and then adding initial vocabulary to see if there is a main effect of the number of words a child knows.

Model fit statistics are calculated at each step as predictors are added to the model. In this study, -2 log likelihood (-2LL, a value which increases as model fit increases) will be used, and a chi-squared test will show whether the increase in -2LL is significant (Singer & Willett, 2003).

<u>Analysis plan</u>

There were three main goals in the analysis plan for this study. The first was to construct models for the four measures of language growth based on time and diagnosis group. Second, we added theoretically important predictors (i.e., from Table 4) to the growth models to create well-fit models. The third goal was to use the growth estimates in those models to predict the V7 outcome measures.

In order to reach the first goal, we examined plots of each of our four measures of language growth to visually assess their overall temporal patterns and group differences. We divided the data by diagnosis group, or by subgroup, depending on the measure in question. (See Figures 2 through 5 for the data.) By examining the raw data, we were able to conjecture whether there were group differences, and whether a linear, quadratic, or cubic line might best be used to model the measure. This step helps avoid overfitting as discussed above. We then created and tested models with visit and diagnosis as predictors. In all cases, we used a random effect on

Participant and visit. In cases where we found quadratic effects on visit, we included those as random effects as well[3]. Including the random effects meant that we took into account individual variation in intercept and slope, as well as acceleration (quadratic effects).

The second goal was to develop the model further to include other theoretically important variables. We used social and computational variables from the initial visit to see if they added significantly to the models. We used both evidence of a significant predictor as well as improvements in model fit to judge whether predictors were important to models. In some cases, we added predictors in interaction with the visit variable, and in others we simply added them as a main effect. A predictor in interaction with visit affects the slope, while a predictor added as a main effect just affects the intercept. Key to this step was also the addition of the overall number of word tokens a child used during each play session, in order to control for a child's talkativeness in the face of his or her other abilities. This allows us to analyze word types and pronouns as simple counts without worrying that they are affected by the child's general level of talk. We added word tokens as an intercept (or main) effect, rather than a slope effect, so that we could simply use it as a control for intercept differences, rather than to look for an effect on slope. After finding a model that was well-fit and included variables that were important to our theoretical questions, we plotted the model's predictions. This revealed the overall shape of the developmental trajectories, and ideally (if the model is a good fit) looked similar to the raw data.

For the third goal we used a technique that combined growth curve analysis with OLS regression. This method is adapted from Rowe and colleagues (2012) and Carlson, Demir, Goldin-Meadow, and Levine (2013). After completing goal 2, we used the full models of our measures of language growth to find individual growth parameters (for each individual child). In

[3] Our analyses revealed no significant effects higher than quadratic, so there will be no further discussion of cubic effects.

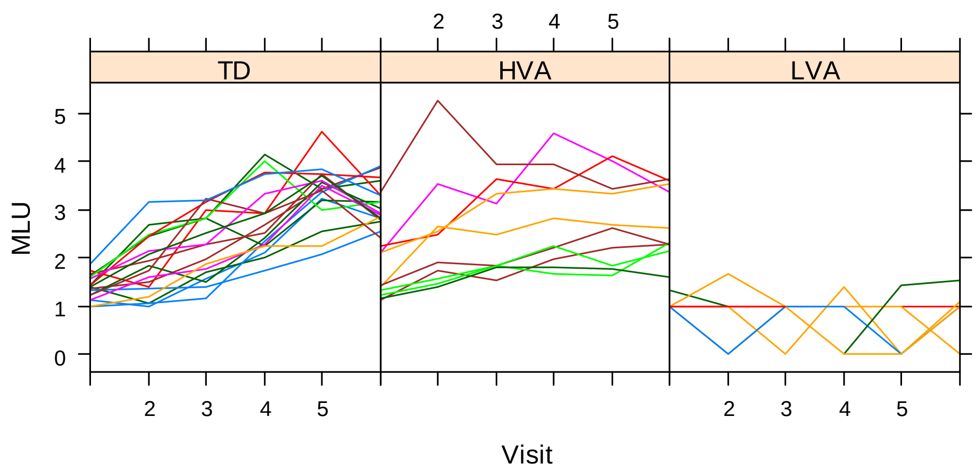
order to do this, we found the coefficient estimated for the intercept, for the slope (the coefficient on visit), and for the quadratic effect when applicable (the coefficient on visit squared). These are like beta weights in OLS regression – they give us an average number that applies equally to all participants. However, in the process of creating our growth curve models, we set the program to calculate the random effect for each of those variables as well. This means that we could then extract those random effects (the zeta values from Figure 1), which are the values for each child that shows how different that individual is from the average value. So, if a child has a random effect of -0.2 on visit, and the model gives a coefficient of 1 on visit, that particular child's visit value is 0.8. We calculated this by-child adjustment for each child, and ended up with a new set of values for each child: the individual growth parameters. The individual growth parameters were then used as data points in new OLS regressions, with visit 7 outcome measures as the dependent variables. This allows us to draw conclusions about whether the *rate of change* predicts later language abilities. For instance, Rowe et al. (2013) used this technique to find that both the linear and quadratic trends in vocabulary growth over time during toddlerhood predicted later vocabulary ability.

3. Results

Goal 1: Create growth models using visit and diagnosis or subgroup for four measures of language growth

Figure 2 shows the raw data plots for the MLU measure, divided by subgroup. Table 6 outlines the growth models for MLU. The columns under Model 1 demonstrate that a quadratic trend best describes the overall trend in MLU over time, regardless of group. Addition of the subgroup variable to create Model 2 significantly improves model fit (demonstrated by increase in -2LL). The significance levels listed under Model 2 reveal the following: the overall intercept is positive and different from zero, the overall linear trend is positive, and the quadratic trend is negative. The HVA group has a significantly different intercept from the TD group[4], while the LVA group does not. There is no significant difference between the linear slope of the HVA group and the TD group, but the LVA group has a significantly shallower slope (because of the negative estimate). The HVA group does not have a significantly different quadratic trend from the TD group, while the LVA group has a significantly less negative quadratic trend than the TD group.

Figure 2. Data plots for mean length of utterance for three subgroups.



[4] In all models, the subgroup and diagnosis variables were designated as factors, with the TD group as zero. This is why all comparisons are made with respect to the TD group.

Table 6. Growth models for mean length of utterance (MLU)

	Model 1: Time				Model 2: Add subgroup			
	Estimate	SE	t-value	p-value	Estimate	SE	t-value	p-value
Intercept	1.40	.10	14.00	<.001**	1.32	.13	10.31	<.001**
Visit	0.45	.09	4.92	<.001**	0.66	.10	6.55	<.001**
Visit^2	-.04	.02	-2.62	.01*	-.05	.02	-2.72	.007*
High-verbal					0.51	.21	2.38	.02*
Low-verbal					-0.26	.24	-1.09	.26
High-verbal x Visit					-0.11	.17	-0.65	.51
Low-verbal x Visit					-0.90	.20	-4.58	<.001**
High-verbal x Visit^2					-.02	.03	-.60	.54
Low-Verbal x Visit^2					0.10	.04	2.48	.01*
-2LL	372.42				306.68			

Figure 3 and Table 7 show the data and growth models for the development of word types[5]. Model 1 demonstrates that the overall trend in word type development fits best with a quadratic trend. Addition of the subgroup variable improves model fit, and all of the estimates

[5] The Word types measure and the Pronoun production measure are count variables that include many zeros. In order to correct for this, we used the poisson distribution rather than a normal distribution, which is somewhat similar to using a log transform, and is appropriate for count data that are not normally distributed (Cameron & Trivedi, 1998).

are significant. This means that both the HVA and LVA groups have different intercepts, linear slopes, and quadratic effects from the TD group. The HVA group has a higher intercept than the TD group while the LVA group has a lower intercept. Both ASD groups have shallower linear slopes but less negative quadratic effects than the TD group.

Figure 4 and Table 8 show the data and growth models for the development of pronoun production. The figure looks somewhat different from the other measures, because there is a line for each type of pronoun (1^{st}, 2^{nd}, and 3^{rd} person) for each child. The different pronoun types will be accounted for in goal 2, when we create a fuller model of pronoun development. The table shows that a quadratic term best fits the overall development of pronoun production. Adding the diagnosis variable improves model fit, and we see that the overall intercept is not different from zero and the overall slope has a significant negative quadratic trend. The ASD group is significantly different from the TD group with regards to the intercept, the linear slope, and the quadratic effect. The ASD group had a higher intercept, a shallower linear slope, and less of a negative quadratic trend.

Figure 5 and Table 9 show the data and growth models for the development of the shape bias comprehension measure. In this case, Model 1 shows that the data are best fit by a linear term, although the slope is not significantly different from zero. For the purpose of replicating and expanding on work that found group differences in shape bias performance (Tek et al., 2008), we added the diagnosis variable and visit variables in interaction with each other, without including main effects. This reveals a significant positive slope for the TD group but not the ASD group.

Figure 3. Data plots for word types for three subgroups.

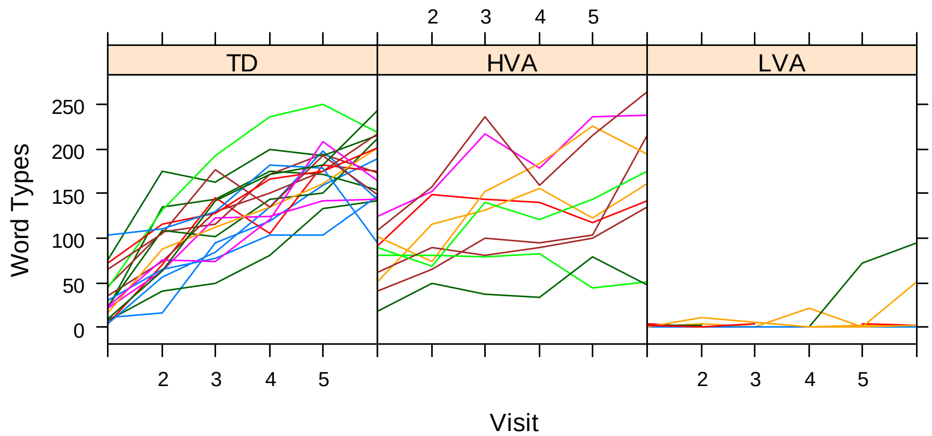


Table 7. Growth models for word types

	Model 1: Time				Model 2: Add subgroup			
	Estimate	SE	z-value[6]	p-value	Estimate	SE	z-value	p-value
Intercept	3.10	.28	11.04	<.001**	3.56	.14	24.61	<.001**
Visit	0.52	.07	7.60	<.001**	0.76	.06	12.13	<.001**
Visit^2	-0.05	.01	-5.31	<.001**	-0.09	.01	-11.31	<.001**
High-verbal					0.73	.24	3.04	.002*
Low-verbal					-3.23	.37	-8.63	<.001**
High-verbal x Visit					-0.52	.10	-5.02	<.001**
Low-verbal x					-0.48	.22	-2.16	.03*

[6] The models that were made using the poisson distribution (i.e. word types and pronouns) resulted in predictors being assigned a z-value rather than a t-value.

Visit					
High-verbal x Visit^2		0.07	.01	5.08	<.001**
Low-verbal x Visit^2		0.09	.03	2.75	.005*
-2LL	1190.78	1084.4			

Figure 4. Data plots for the raw number of pronouns produced by the two diagnosis groups. There are three lines per child in these plots: one each for 1st, 2nd, and 3rd person pronouns.

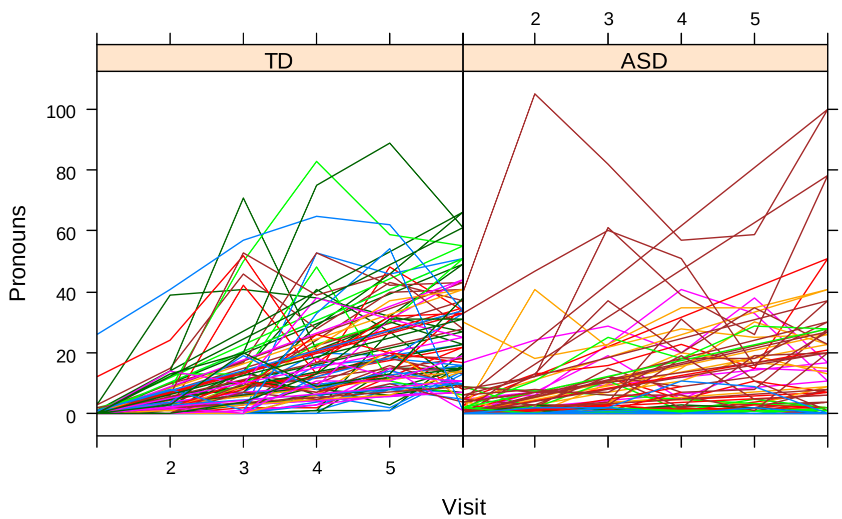


Table 8. Growth models for pronoun production.

	Model 1: Time				Model 2: Add Diagnosis			
	Estimate	SE	z-value	p-value	Estimate	SE	z-value	p-value
Intercept	-0.05	.32	-.17	.87	-0.61	.40	-1.52	.13
Visit	1.28	.16	8.21	<.001**	1.76	.15	11.73	<.001**
Visit^2	-0.15	.02	-7.23	<.001**	-0.21	.02	-10.01	<.001**
ASD group					1.28	.59	2.18	.03*
ASD group x Visit					-1.10	.22	-4.95	<.001**
ASD group x Visit^2					0.13	.03	4.21	<.001**
-2 LL	5364.4				5339.0			

Figure 5. Data plots for shape bias comprehension looking time measure. The y-axis displays the percent of time spent looking at the target object during the Name trial minus the percent of time spent looking during the No Name trial. Higher positive values indicate a stronger use of the shape bias.

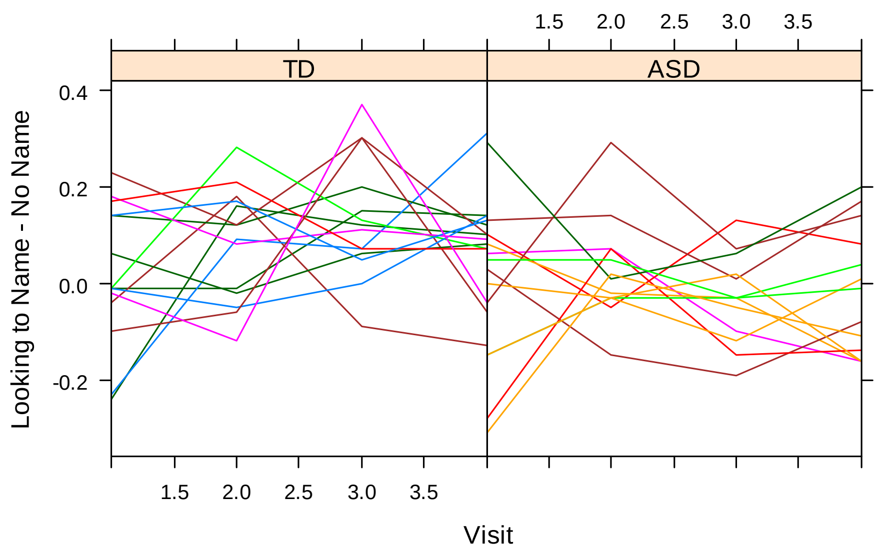


Table 9. Growth models for shape bias comprehension

	Model 1: Time				Model 2: Add diagnosis			
	Estimate	SE	t-value	p-value	Estimate	SE	t-value	p-value
Intercept	.02	.02	.95	.34	.02	.02	.95	n.s.[7]
Visit	.01	.01	.85	.39				
Visit x TD group					.03	.01	2.52	<.05*

[7] The program that gives specific p-values will not run on models that do not include main effects, so these t-values are interpreted by comparing to the critical t-value of 1.96.

Visit x ASD group		-.01	.01	-1.10	n.s.
-2 LL	-134.94	-145.6			

To summarize, the shape of all three production variables is quadratic, with a positive coefficient on the linear effect and a negative coefficient on the quadratic effect. This means that the progression of these three measures is positive but that the rate of increase slows over time; there is a deceleration in the progression of MLU, word types, and pronoun production. The shape bias comprehension increases linearly, without any change in speed over time. All four measures have significant subgroup or diagnosis effects, meaning that there are differences between the way that TD children and children with ASD are progressing over time.

Goal 2: Develop and plot growth curves including important predictors.

In goal 1, we created a model of MLU using linear and quadratic effects on visit and subgroup as predictors. In this step, we added joint attention variables and early language variables to see whether they had any effect on the models. We were also interested in controlling for the general talkativeness of each child, so we included the overall number of word tokens the child used at each visit as a predictor. This is a way of ensuring that the relationships between the predictors and the outcome variables weren't based on sheer number of words that a child produced; using count variables could lead to an extremely talkative child using many pronouns, for instance, but not significantly more per word than a child who is more reticent. Using word tokens as a control helped us avoid spurious findings.[8] Table 10 displays the final model that we developed including the important variables, and the effects are explained below.

[8] Another way of controlling for overall word tokens would have been to use a ratio of types to tokens; we chose not to do this because of the difficulty using ratios as a dependent measure in regression analyses. When these

There is an overall linear slope with a negative quadratic (non-significant) trend. The HVA group has a significantly higher intercept from the TD group, and the LVA group has a significantly shallower linear slope from the TD group. The rate of deceleration (the quadratic trend) is not different between the groups. The number of word tokens a child uses has a main effect on growth, meaning that children who use more word tokens have a higher intercept than those who use fewer, regardless of diagnosis. The child's CDI score at V1 also has a significant main effect, meaning that children with higher vocabularies at the first visit have a higher MLU intercept. These last two points are perhaps not surprising, but they do mean that the relationships between subgroup and visit hold over and above the effects of talkativeness and initial vocabulary. Adding joint attention measures, Vineland Social, Word Order, and Noun Bias measures did not improve the model fit or result in a significant predictor, so they are not included in this model. Figure 6 displays the model predictions for this model.

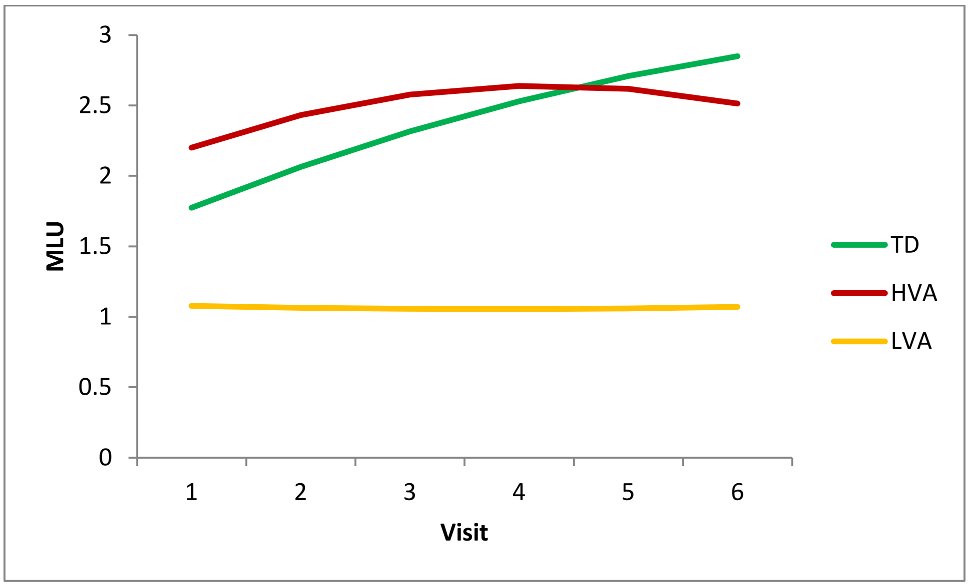
Table 10. Final model for MLU, including subgroup and Word Tokens.

	Estimate	SE	t-value	p-value
Intercept	0.95	0.09	11.08	<.001**
Visit	0.32	0.07	4.41	<.001**
Visit^2	-0.02	0.01	-1.58	0.116
HVA group intercept	0.33	0.13	2.52	0.013*
HVA x Visit (slope)	-0.04	0.12	-0.37	0.712
HVA x Visit^2 (acceleration)	-0.02	0.02	-0.99	0.319
LVA group intercept	0.002	0.13	0.02	0.985

measures are transformed into ratios their distributions are dramatically skewed, as well as being bounded between 0 and 1.

LVA x Visit (slope)	-0.34	0.12	-2.81	0.006*
LVA x Visit^2 (acceleration)	0.02	0.02	1.04	0.299
CDI V1	0.57	0.18	3.20	.003*
Word Tokens	0.001	0.0001	9.65	<.001**

Figure 6. Model predictions for MLU model listed in Table 12.



We followed the same steps for the word types model, adding in joint attention and word token variables to see how the model fit changed. Table 11 displays the final model that we developed. In this case, joint attention as measured by our RJA change variable did contribute to the model. A description of the significant predictors follows: There is an overall positive linear slope as well as a negative quadratic effect. The HVA group has a significantly higher intercept, less steep linear slope, and less of a negative quadratic effect than the TD group. The LVA group has a significantly lower intercept from the TD group. The RJA change variable has a significant effect on the overall slope, meaning that children with more positive RJA change values have steeper word types slopes. The RJA change variable had significantly more of an impact in the LVA group, with a steeper slope for that group in interaction with RJA change than in the TD group. The word tokens measure had a significant impact on intercept, with children who spoke more having a higher intercept. Including this variable also means that all the relationships listed above hold over and above general talkativeness. Figure 7a displays the model predictions for children who have an average value on the RJA change variable. In Figure 7b, we took two different values of RJA change for each group – the 1st quartile and 3rd quartile values – and plotted the model predictions for a low and high value of RJA change.

Table 11. Final growth model of word types.

	Estimate	SE	z-value	p-value
Intercept	3.51	0.13	26.24	<.001**
Visit	0.52	0.06	8.62	<.001**
Visit^2	-0.06	0.01	-6.34	<.001**
HVA (main)	0.59	0.28	2.13	0.03*
HVA x Visit (slope)	-0.36	0.12	-2.98	0.002*
HVA x Visit^2 (acceleration)	0.05	0.02	2.29	0.02*
LVA (main)	-3.02	0.37	-8.21	<.001**
LVA x Visit (slope)	-0.26	0.24	-1.08	0.28
LVA x Visit^2 (acceleration)	0.06	0.04	1.50	0.13
RJAchange (main)	-0.007	0.005	-1.51	0.13
RJAchange x Visit (slope)	0.004	0.002	1.95	0.05*
RJAchange x Visit^2 (acceleration)	-0.0004	0.004	-1.38	0.17
HVA x RJAchange (main)	0.01	0.02	0.67	0.50
HVA x RJAchange x Visit (slope)	-0.002	0.01	-0.25	0.80
HVA x RJAchange x Visit^2 (acceleration)	0.0002	0.001	0.20	0.84
LVA x RJAchange (main)	-0.09	0.04	-2.67	0.01*
LVA x RJAchange x Visit (slope)	0.05	0.02	2.46	0.01*
LVA x RJAchange x Visit^2 (acceleration)	-0.005	0.003	-1.83	0.07
Word Tokens (main)	0.0008	0.0001	15.68	<.001**

Figure 7a. Model predictions for word types model listed in Table 13, with average value for RJAchange.

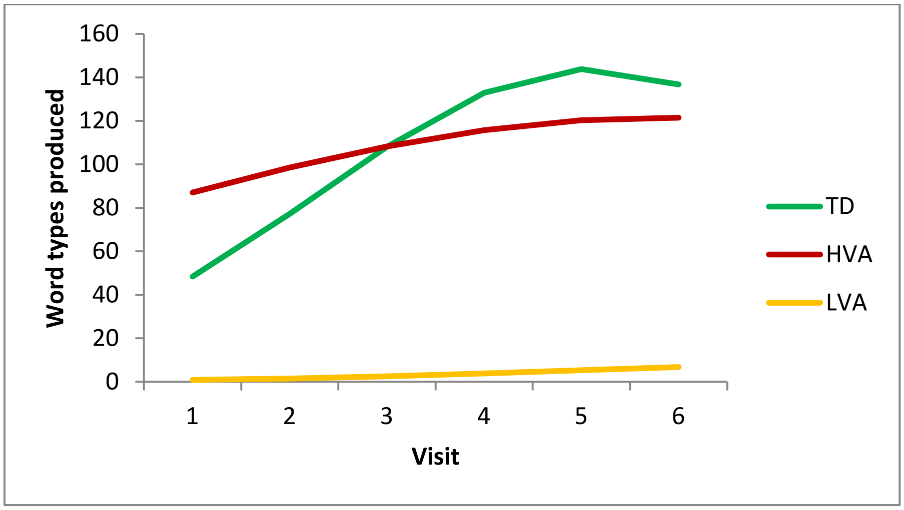
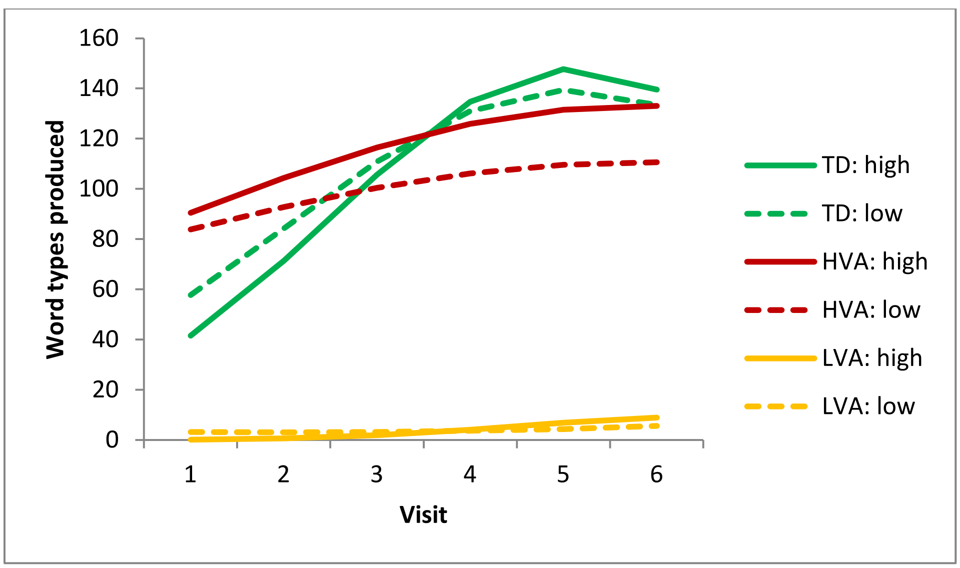


Figure 7b. Model predictions for word types model listed in Table 10, with high and low values for RJAchange.



The pronoun model is more complicated than the previous two, because there are three different types of pronouns being modeled. Table 12 displays the final model, which contains visit, diagnosis, pronoun type, RJA change, and word tokens and noun bias as predictors. The pronoun types were coded such that 1^{st} person pronouns were the comparison point (this is because 1st person pronouns are generally acquired first; Charney, 1980), so all the 2nd and 3rd person pronoun effects are relative to the 1st person effects. A description of the significant predictors follows: There is an overall positive linear slope with a negative quadratic effect. The 2nd person pronouns (regardless of diagnosis group) have a lower intercept than the 1st person pronouns, and a steeper linear slope. There is a lower intercept for 3rd person pronouns (regardless of diagnosis group), as well as a steeper slope and a more negative quadratic effect. The ASD group has a higher intercept than the TD group for pronouns overall, a shallower slope, and a less negative quadratic effect. The ASD group has a steeper slope for 3rd person pronouns and a more negative quadratic effect. The RJA change variable interacts with visit to show an overall steeper slope and more negative quadratic effect overall; none of the RJA change interactions with diagnosis were significant, meaning that the effect of the RJA change variable was similar for the two groups. Finally, the word tokens measure and the noun bias looking time measure had significant effects on the intercept, so that children who used more words had a higher intercept, and children who had less of a noun bias at V1 had a higher pronoun intercept. Figure 8a shows the model predictions for both groups and all three pronoun types using the average RJA change value for each group. Figure 8b displays only the 1st person pronouns for high (3rd quartile) and low (1st quartile) values for RJA change to demonstrate the effects of joint attention on pronoun production. A higher score on noun bias simply shifted the intercept down, rather than changing the trajectories of development.

Table 12. Final growth model for pronoun production.

	Estimate	SE	z-value	p-value
Intercept	0.45	0.32	1.40	0.16
Visit	1.17	0.13	8.86	<.001**

$Visit^2$	-0.15	0.02	-8.74	<.001**
2nd person	-2.62	0.28	-9.46	<.001**
2nd person x Visit (slope)	0.38	0.17	2.24	0.02*
2nd person x $Visit^2$ (acceleration)	-0.003	0.02	-0.11	0.91
3rd person	-2.47	0.24	-10.13	<.001**
3rd person x Visit (slope)	0.72	0.15	4.74	<.001**
3rd person x $Visit^2$ (acceleration)	-0.08	0.02	-3.70	<.001**
ASD	1.47	0.68	2.17	0.03*
ASD x Visit (slope)	-1.01	0.26	-3.84	<.001**
ASD x $Visit^2$ (acceleration)	0.14	0.03	4.20	<.001**
ASD x 2nd person	0.47	0.35	1.32	0.24
ASD x 2nd person x Visit (slope)	0.10	0.24	0.43	0.67
ASD x 2nd person x $Visit^2$ (acceleration)	-0.05	0.04	-1.28	0.18
ASD x 3rd person	-0.54	0.37	-1.46	0.14
ASD x 3rd person x Visit (slope)	0.59	0.25	2.38	0.01*
ASD x 3rd person x $Visit^2$ (acceleration)	-0.11	0.04	-2.84	0.004*
RJAchange	-0.02	0.01	-1.70	0.09^+
RJAchange x Visit (slope)	0.01	0.004	2.16	0.03*
RJAchange x $Visit^2$ (acceleration)	-0.001	0.001	-2.30	0.02*
ASD x RJAchange	0.04	0.04	1.20	0.23
ASD x RJAchange x Visit (slope)	-0.02	0.02	-1.08	0.28
ASD x RJAchange x $Visit^2$ (acceleration)	0.003	0.002	1.68	0.09^+
Noun Bias % looking time	-1.05	0.455	-2.31	0.02*

Word Tokens	0.001	<.001	13.87	<.001**

Figure 8a. Model predictions for the pronoun production model listed in Table 11, with average values for RJAchange.

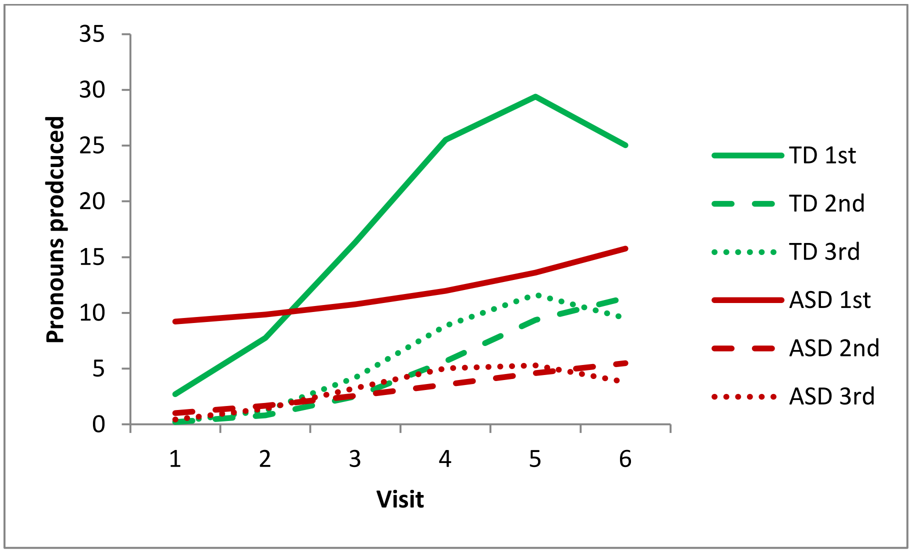


Figure 8b. Model predictions for 1st person pronouns, with high and low values for RJAchange.

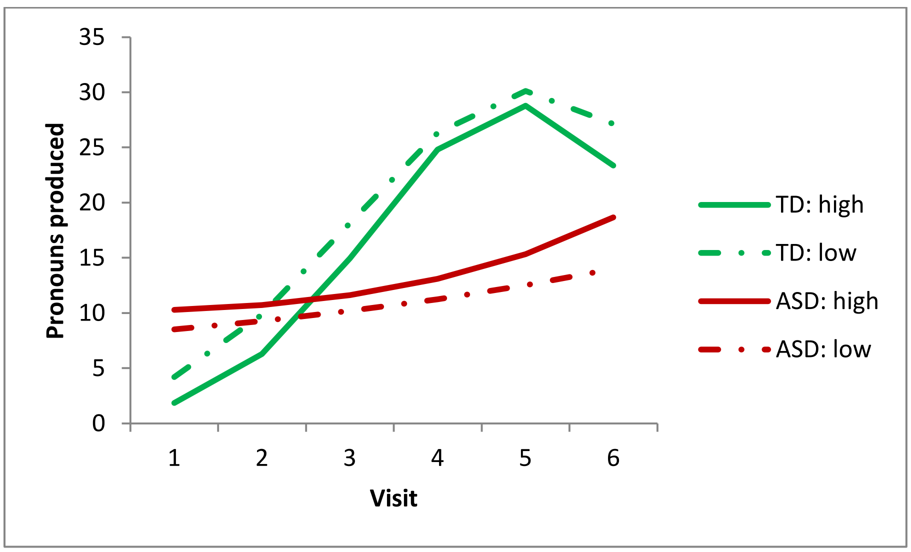


Table 13 displays the growth model for our shape bias comprehension measure. Tek et al. (2008) used growth curves to analyze the same data and found that children with ASD did not use the shape bias (i.e. did not get positive scores on this measure), and moreover they did not increase in their scores over time, while the TD children did increase. The present analysis included effects of joint attention to enhance our understanding of the shape bias difference in ASD and TD children. There is an overall negative intercept and positive linear slope. The ASD children have a significantly shallower slope than the TD children. The Initial RJA measure significantly contributed to this model, with a positive effect on intercept and slope. This means that higher levels of initial RJA lead to more positive values of the shape bias comprehension measure. Figure 9a displays the measure for both groups with an average level of initial RJA, and Figure 9b displays the effect of high and low levels of initial RJA for both groups. Although it appears that the two groups respond differently to differing levels of initial RJA, the data do not support a model that includes an interaction between diagnosis and initial RJA.

Table 13. Final growth model for shape bias comprehension measure.

	Estimate	SE	t-value	p-value
Intercept	-.120	.043	-2.80	<.01*
Visit	.076	.025	3.01	<.01*
ASD x Visit (slope)	-.042	.013	-3.22	<.01*
Initial RJA	<.01	<.01	3.71	<.01*
Initial RJA x Visit (slope)	<.01	<.1	-2.22	.02*

Figure 9a. Shape bias growth calculated using an average score on initial RJA

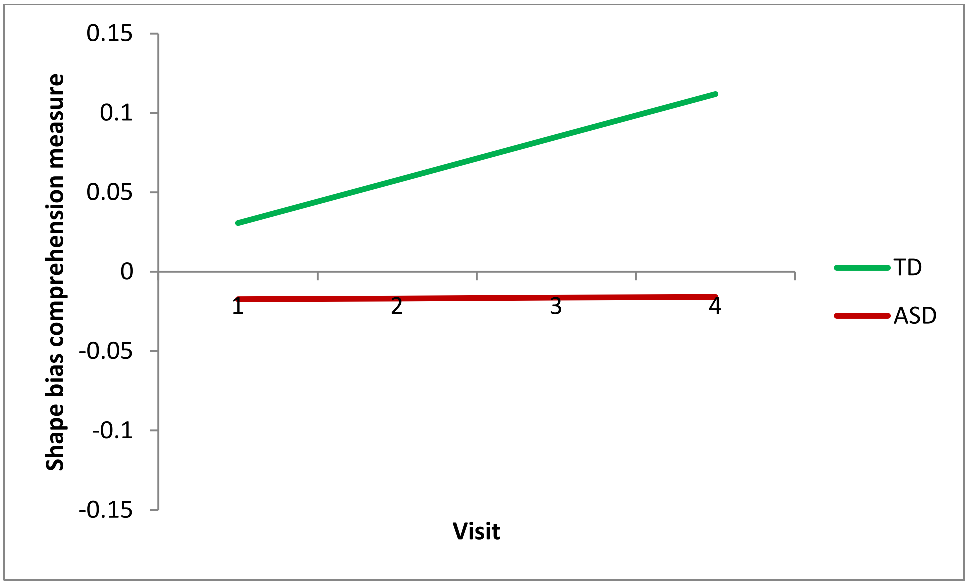
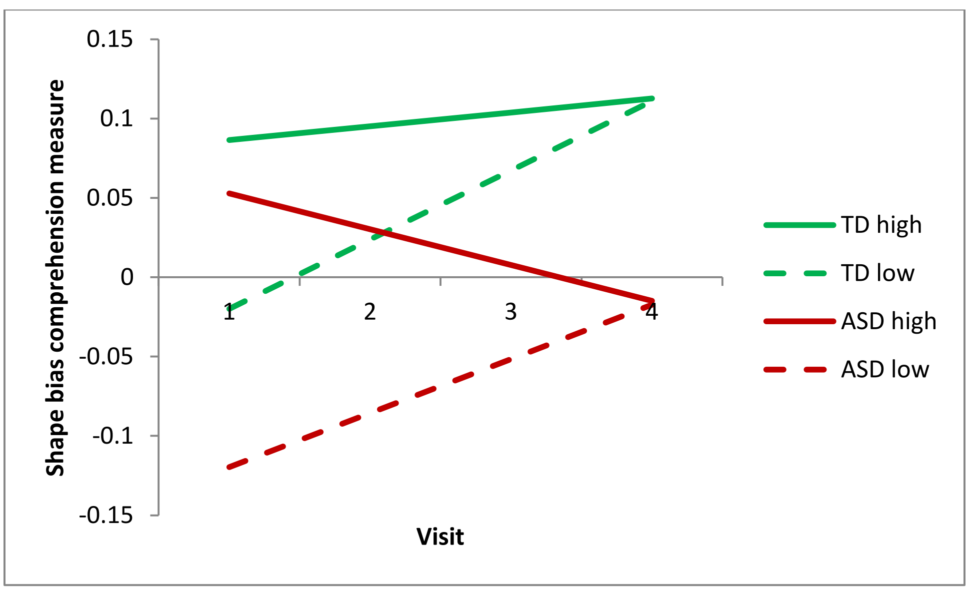


Figure 9b. Shape bias growth calculated using a high and low score on RJA



To summarize, the one grammatical measure, MLU, showed no influence of joint attention measures and only an influence of initial vocabulary in the computational measures. Two of the lexical measures, word types and pronouns, showed positive effects of the RJA change measure on development overall, with the effect being different across diagnosis groups on word types but not on pronouns. All three production measures had a significant positive coefficient on word tokens, meaning that children who talked more had higher scores on all three measures. For MLU, in fact, adding this variable washed out the difference between the TD and HVA groups that was revealed under goal 1, suggesting that when overall talkativeness is taken into account there is no significant difference in those two groups on MLU trajectories. The shape bias measure showed group differences on slope, and an effect of Initial RJA, but no group differences in the effect of joint attention.

Goal 3: Predict V7 measures from growth parameters

We chose outcome measures of interest from our V7 in order to see whether the rate of early language development could predict later language and cognitive outcomes. Each outcome measure was used as a dependent variable (see Table 5). The independent measures were: subgroup and the individual growth parameters for MLU; subgroup and the individual growth parameters for word types; diagnosis and the individual growth parameters for pronoun production; and diagnosis and the individual growth parameters for shape bias comprehension, respectively.

None of the individual growth parameter models predicted scores on the Vineland Social, the IQ as measured by the DAS, the TACL syntax measure, the wugs task, the mental verbs task, or the theory of mind task.

The TACL vocabulary scale was predicted by the model including word types, and the TACL morphology scale was predicted by two different models: one involving the MLU growth parameters, and one involving the word types growth parameters. See Table 14 for the model of TACL vocabulary scores. This means that children who have higher intercepts and steeper slopes in terms of word type development are those children who tended to have higher scores on the TACL vocabulary test at V7. Table 15 displays two different models that predict the TACL morphology scores at V7. The left panel shows a model that includes a marginal effect of MLU slope, meaning that the children with steeper MLU trajectories have a trend towards higher TACL morphology scores. The right panel shows a model that includes a significant effect of the quadratic effect of word types, so that children with a stronger negative quadratic trend for word type development are those with higher TACL morphology scores. Adjusted R-square measures demonstrate that the model including word types growth explains more variance than the model including MLU growth.

The last type of regression model we did examined the possibility that a child's diagnosis category could be predicted by the individual language growth parameters. In order to do this, we recalculated the growth parameters using growth curve models that did *not* include any diagnosis or subgroup variable, and thus treated all the children as if they were in one group. All other variables remained the same. We then did logistic regression with the outcome as a binary variable: either diagnosed as having ASD or not. The only model that showed any relationships of interest included the Shape Bias growth parameter, and it is shown in Table 16. This shows a marginally significant trend towards children with lower intercepts on the shape bias measure as more likely to be in the autism group.

Table 14. Regression model predicting TACL vocabulary scores at Visit 7.

	Beta	t-value	p-value
Subgroup	-.879	-10.76	<.001
Word Types Intercept	.417	1.92	.07^{+}
Word Types Slope	.410	1.90	.07^{+}

Table 15. Two different regression models predicting TACL morphology scores at Visit 7.

	Beta	t-value	p-value		Beta	t-value	p-value
Subgroup	-.889	-10.05	<.001**	Subgroup	-.841	-10.13	<.001**
MLU Intercept	-.832	-1.62	.118	Word types intercept	-.116	-.35	.73
MLU Slope	.913	1.77	.09^{+}	Word types slope	.344	1.55	.13
				Word types acceleration (quad)	.612	2.13	.04*
Adjusted R-squared		*.77*				*.80*	

Table 16. Logistic regression model predicting autism diagnosis at Visit 7.

	Beta	SE	p-value
Shape Bias intercept	-67.48	41.19	.10^{+}
Shape Bias slope	-111.45	235.24	.64

The regressions in Tables 14 and 15 all have a significant or trending-toward-significant predictor on a slope parameter variable. This means that the growth (change over time) of the predictor (MLU or word types) is a better predictor of the outcome variable than the initial level of MLU or word types. This supports the idea that change over time gives us more information than a simple beginning time point. The regression in Table 16 has only an intercept effect, but further work on this question could include more participants to see if there is an effect of shape bias growth in predicting a diagnosis of ASD.

4. Discussion

This dissertation is a study of social and computational contributions to specific language measures in children with ASD and TD children. Many major theories of language development hinge on whether children primarily use social or computational skill to develop language, but few have tried to directly compare the contributions of these abilities to language development, and to our knowledge none have done so in a longitudinal and detailed way.

Our analyses for goal 1 revealed that all four measures of language growth increased over time for all groups with two exceptions: MLU in the LVA group, and shape bias in the ASD group. Goal 2 analyses revealed the influences of the computational and social variables. MLU showed an effect of initial vocabulary overall (a main effect) and of word tokens, but no effect of other computational measures or of joint attention. Word types growth is steeper for HVA children than TD children, and slower for LVA children than for TD children. Word types growth showed an effect of RJA change on growth, which was stronger in the LVA group than the TD group, as well as an overall effect of word tokens. Pronoun production showed that the children with ASD produced fewer pronouns overall and increased more slowly in their production than the TD children, and that RJA change had a similarly positive influence on both groups of children. In addition, noun bias had a negative overall influence on pronoun production, while word tokens had a positive influence. Shape bias growth was different between groups; the ASD group increase was shallower than the TD group, and initial joint attention had a positive effect both overall and on the slope of shape bias for both groups. Goal 3 analyses revealed a trend that steeper word types growth predicted higher TACL vocabulary scores at V7, and that both MLU linear growth and word types quadratic (decelerating) growth predicted TACL morphology scores independently. Shape bias growth trended towards a prediction of diagnostic category (ASD versus TD).

We will discuss these findings in four parts below: specific influences of computational factors, specific influences of social factors, diagnostic group differences, and meaning of the higher order shapes of growth.

Contributions of computational and social measures. Computational measures only showed overall, or main, effects on our measures of language growth. This means that computational measures did not influence the slope of the measures, but just the overall level of those measures. On a graph, these effects would be indicated by an intercept shift, rather than a different shape or rate of change. Additionally, there were no computational measures that differed in their influence by diagnostic group; any of the computational measures that had an influence did so similarly across groups. Initial vocabulary level had a positive effect on MLU, meaning that children with higher initial vocabularies had higher MLU overall. Noun bias had a negative main effect on pronoun production. Children with higher noun bias scores produced fewer pronouns overall. Word tokens had a main effect on all three production measures, but again this was being used as a control in these analyses rather than as a measure of particular theoretical interest. Other measures did not show significant influence of early computational measures.

Joint attention change had an influence on slope for both the word types and the pronouns measures. Higher joint attention change scores (made by finding the difference between each child's early joint attention skills across time compared to the mean) increased the slope of word types and pronouns. Initial joint attention influenced overall shape bias scores and the slope of shape bias scores, meaning that children who had higher levels of early joint attention had higher scores on the shape bias measure and had steeper increases in shape bias scores. Whether these effects differed by diagnosis groups will be discussed immediately below.

These analyses add interesting fuel to the debate between social and computational theories of language development. As described in the introduction, most of the data linking joint attention to language learning is about lexical learning. The change in MLU of typical and atypical children did not depend on their social abilities as measured in this study. Since all of our other measures were, in some way, lexically-based, and showed an effect of joint attention, this is evidence for the idea that lexical learning may be more influenced by joint attention than grammatical learning. Joint attention measures predicted lexical growth, and vocabulary (measured by the CDI) contributed to MLU, but there was no direct influence from social measures to grammatical measures. The cues to word reference are clearly present in joint attention episodes; a parent can point directly at an object while naming it, but it seems that the cues to stringing those words together are not present in social engagement.

Different researchers have proposed different roles for joint attention depending on children's age and language level. Hoff and Naigles (2002) suggest that children use joint attention (with lots of help from parents) early on to grasp the meanings of words and then begin using more computational abilities to develop further vocabulary and grammar. Hollich et al. (2000) showed evidence of children using joint attention (in an experimental setting, without parental input) *later* in the word learning process rather than earlier. The current findings show that those children who have higher RJA change scores have steeper increases in their word types, unlike the findings of Hoff and Naigles that joint attention become less important in later word learning. However, our measure is of *response* to joint attention, and Hoff and Naigles point out that the children in their study got a lot of assistance from their parents in their joint attention episode, which may be something more like initiation of joint attention (with the initiation being done by the child). Perhaps initiation is important early on, and then episodes of joint attention that the child enters by responding get increasingly important in word learning. This would also explain the results of Hollich et al. that have children responding to an experimenter's bid for attention, that show later use of joint attention for word learning. Interestingly, our study showed children with LVA who had higher RJA change scores had even steeper increases in word types growth than the TD children. Perhaps those children with LVA are delayed to the point where their increases in response to joint attention are that much more salient to their word production abilities.

The pronoun measure was the only one that revealed an influence of a social and a computational measure. Noun bias had a negative main effect on pronoun production, such that children who had higher scores on the noun bias measure showed lower overall numbers of pronouns. This could mean that children who are more able to learn nouns are less likely to use pronouns. Jordan (1989) and Lee, Hobson, and Chiat (1994) found that children with ASD were more likely to use names and nouns instead of pronouns, and while our noun bias finding was the same across diagnostic groups, it could be that children who are more engaged in learning nouns simply use those words rather than pronouns. RJA change had a positive influence on the slope of the pronoun measure. Engagement in joint attention serves as a rich context for using pronouns, because the referent of a pronoun is easily indicated by points or eye gazes. In this study, pronouns were considered to be lexical because the measure was simply the number produced. However, if we considered the child's use of specific pronouns with regards to number, case, and person, and saw whether they were using the correct pronoun in each instance, it might be considered a grammatical measure. Altering the pronoun measure could then lead us to different and complementary conclusions about pronoun development.

Differences in diagnostic groups. At least some of the children with ASD had flatter trajectories on all of our measures. For MLU, children with LVA showed no significant improvement over time, while children with HVA were on par with TD children. This corroborates the findings of Tek et al. (2013), whose study showed similar paths of development for TD and HVA children on many measures. With regard to word types, both ASD groups had slower increases and less deceleration (i.e., did not level off as much at the later visits) than the TD group. The children with ASD had shallower slopes on pronoun production, producing fewer pronouns overall and increasing more slowly across time. The children with ASD also had a shallower slope on the measure of shape bias comprehension. In short, the only case in which the TD group and an ASD group did not differ was the HVA children on MLU. Tek et al. (2013) found that children with HVA had a significantly shallower slope than TD children for MLU. This discrepancy can be explained by the fact that the current study used the number of word tokens as a predictor in order to control for children's talkativeness. Before word tokens were included in the model (under goal 1 in Results; see Table 6) the HVA group did not have a significantly different slope from the TD children. Since the TD children were, on average, more talkative, once the number of word token was controlled for this, the significant slope difference disappeared.

There was one case in which joint attention had different effects depending on diagnosis group, but no cases where computational effect had different effects on diagnosis groups. Joint attention had a stronger influence on children with LVA than TD children in terms of their word types growth. LVA children with higher scores on the joint attention change measure had lower intercepts than TD children (not surprising since this group had significantly lower intercepts on many production measures) but steeper slopes; their high scores on joint attention change increased their slope of pronoun production. This may relate to the point made earlier about whether joint attention has different effects depending on the language level of the child. Children with LVA started the study with very few words and an MLU of about one, so perhaps their very early stage in language development meant that joint attention served them well, while the HVA and TD children had already moved on to using computational abilities (albeit none that were captured in this study), similar to the suggestions based on Hoff and Naigles (2002). This is not to say that the higher-functioning children were *not* using joint attention; indeed their trajectories were steeper when a higher level of joint attention was modeled. However, the difference that an increase in joint attention made was significantly larger for children with LVA.

Higher order shapes in language growth. Our production measures were all modeled with quadratic effects, which led to the appearance of interesting group differences in the word types and pronouns measures. TD children had negative quadratic effects, meaning that there was a deceleration in their progression. However, children with ASD had less of a quadratic effect on both measures, meaning that they kept up the pace of increase more than the TD children did.

One reason for this could be a general delay in the development of language in the children with ASD; perhaps they will also taper off in their word types and pronoun development, but it will happen at a later time than was captured in this dataset. While some data have supported the idea that children with ASD have patterns of language ability that deviate from the norm (Eigsti & Bennetto, 2009; Boucher, 2012), others suggest that underlying processes are similar but simply delayed (Goodwin et al., 2012; Swensen et al., 2007). The current study certainly supports the idea that there is a delay in the development of language in ASD, with the children with ASD being chronologically older than the TD children when matched on language at visit 1. Another possibility for the stronger negative quadratic trend in the TD children is simply methodological: the children played with the same set of toys at each visit, and it is possible that the TD children were more easily bored by the toys and produced fewer new word types by the last visit because they were talking about the same situations and games as in previous visits. Future analyses could determine this by looking at the exact words that the children are using, and trying to see whether they had hit a ceiling in terms of the different words one is likely to use given the situation.

Predicting V7 measures from early language growth

Few V7 measures were modeled by the individual growth parameters. Steeper word types growth predicted higher TACL vocabulary scores, which showed that growth in early vocabulary ability is a good indicator of vocabulary two years later. This corroborates the findings of Rowe et al. (2012) who found that early vocabulary growth was a better predictor of later vocabulary than a simple initial measure. This is also true in our analysis, since the word types *growth* parameter is significant over and above the *intercept* parameter. TACL morphology scores were predicted by growth in MLU, meaning that those children who had faster increases in MLU had higher morphology scores. This points to a relationship between grammatical and morphological learning, and indicates positive results for those children who increase quickly in their grammatical complexity. TACL morphology scores were also predicted by deceleration in word types growth (i.e. the negative quadratic). This suggests that those children whose word type growth tapered off faster had higher morphology scores later. The most likely explanation for this finding is methodological. As mentioned earlier, however, the word types measure may decelerate more among those children who are less interested in the repeated play session. These children may be more advanced in their language and cognitive skills overall and therefore could decrease in word types due to lack of interest, but also have high morphology scores. However, this leaves open the question of why word types deceleration does not predict other abilities such as the wug task or mental verb comprehension.

There are several reasons that the other individual growth parameters may not have been predictive of later outcomes. The first is a simple methodological issue: we had relatively small sample sizes, particularly for the children with ASD at V7. Adding more children (who are currently being run and analyzed) could increase the number of different predictive models, as well as strengthen the reported models, many of which include trends rather than significant effects. Another more complex possibility is that we did not choose the relevant predictor variables that feed into later linguistic and cognitive skill. We were using the four measures of language growth to try and predict other quite specific skills. For comparison, Rowe et al. (2012) used vocabulary growth to predict later vocabulary size, and Carlson et al. (2013) used MLU growth to predict later standardized measure scores of language. In that respect, the current results are in line with those previous studies, because MLU growth and word types growth predicted later standardized measure scores. It could be that the abilities measured by the theory of mind task, mental verb comprehension task, the Wug task, and the DAS (IQ test) have much more contributing to them than any of the single growth curves we used.

The current analyses went a step beyond any studies we know of to see if growth in any of these measures could predict whether a child was diagnosed with ASD. When the individual growth parameters models were recalculated without including the diagnostic categories, the growth of shape bias was the only variable that trended towards a prediction of diagnostic category. Further work with more participants should explore this potential to see if it is a significant finding. If so, the growth of shape bias ability may prove itself to be an early indicator of ASD. It could be that children are using different lexical processes to assign words to objects, and that this relates to their delayed vocabulary growth. Alternatively, this could be support for a growing number of studies that are finding very early differences in perceptual patterns in children with ASD. For instance, Chawarska & Shic (2009) and Tenenbaum, Shah, Sobel, & Malle (2012) find that children who later develop autism are more likely to show atypical scanning when looking at faces. Our particular measure of shape bias was conceptual, that is, we tested whether children had abstracted a pattern out of their input and begun to assume that words for objects refer to their shapes. It did not test how children were visually exploring the object. It could be that children are attending to the object atypically and this is feeding into their lack of shape bias use. Further work could use eye tracking to see how children are exploring the objects and consider perceptual versus conceptual aspects of the shape bias.

Future work should include more children in each diagnostic group. Although this study has a relatively large sample size for a longitudinal study with a special population, sample size is still a limitation, particularly when the group with ASD is split into high- and low-verbal children, and particularly with the shape bias measure, which had a smaller sample size due to data availability. Future analyses of these measures will include more children in all groups (which are currently being collected and coded), and may prove particularly important to the V7 analyses, since there was some attrition at the outcome visit.

Another point for future work pertains to the definition of joint attention. Some researchers have brought up the idea that children with ASD may use different social cues than TD children when they are interacting with communication partners (Akhtar & Gernsbacher, 2008; Gernsbacher, Stevenson, Khandakar, & Goldsmith, 2008). Future work could include more complex measures of social abilities; this might give insight into aspects of social interactions that children with ASD are more attuned to than joint attention. For instance, according to Gernsbacher et al., children with ASD covertly attend to eye gaze and understand intended actions in ways that may not be coded by the current definitions of joint attention behaviors. If that is the case, the RJA measures used in this study may under-represent the shared attention between children with ASD and their parents. Perhaps joint attention is truly working differently across diagnostic groups for measures like pronoun production and shape bias, while this study shows them having similar effects across groups.

This data in this dissertation reveal interesting effects of social and computational abilities on language development. Computational abilities influence TD children and those with ASD with regards to their growth in early grammar abilities and pronoun production, and social abilities influence word type increases, pronoun production, and comprehension of shape bias. Children with ASD were demonstrated to be delayed in their production of increasing number of word types and pronouns, and evidence for a real difference in lexical processing was demonstrated through a lack of shape bias use.

References

Abecassis, M., Sera, M. D., Yonas, A., & Schwade, J. (2001). What's in a shape? Children represent shape variability differently than adults when naming objects. *Journal of Experimental Child Psychology, 78*, 213-239.

Adolph, K. E., Robinson, S. R., Young, J. W., & Gill-Alvarez, F. (2008). What is the shape of developmental change? *Psychological Review, 115*, 527-543.

Akhtar, N. (2004). Nativist versus constructivist goals in studying child language. *Journal of Child Language, 31*, 459-462.

Akhtar, N., Dunham, F., & Dunham, P. J. (1991). Directive interactions and early vocabulary development: The role of joint attentional focus. *Journal of Child Language, 18*, 41-49.

Akhtar, N., & Gernsbacher, M. A. (2008). On privileging the role of gaze in infant social cognition. *Child Development Perspectives, 2*, 59-65.

Akhtar, N. & Tomasello, M. (2000). The social nature of words and word learning. In R. M. Golinkoff, K. Hirsh-Pasek, L. Bloom, L. B. Smith, A. Woodward, N. Akhtar, M. Tomasello, & G. Hollich, (Eds.), *Becoming a word learner: A debate on lexical acquisition* (pp. 115-135). New York: Oxford University Press.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders (Vol. IV-TR)*. Washington, DC: American Psychological Association.

Anderson, D. K., Lord, C., Risi, S., DiLavore, P. S., Shulman, C., Thurm, A., Welch, K. & Pickles, A. (2007). Patterns of growth in verbal abilities among children with autism spectrum disorder. *Journal of Consulting and Clinical Psychology, 75*, 594-604.

Anglin, J. M. (1993). Knowing versus learning words. *Monographs of the Society for Research in Child Development*, *58*(10), 176-186.

Babyak, M. A. (2004). What you see may not be what you get: A brief, non-technical introduction to overfitting in regression-type models. *Psychosomatic Medicine, 66*, 411-421.

Baldwin, D. (1993). Early referential understanding: Infants' ability to recognize referential acts for what they are. *Developmental Psychology, 29*, 832-843.

Baldwin, D., Markman, E., Bill, B., Desjardins, R. N., Irwin, J. M., Tidball, G. (1996). Infants' reliance on social criterion for establishing word-object relations. *Child Development, 67*, 3135-3153.

Baldwin, D. & Meyer, M. (2007). How inherently social is language? In E. Hoff & M. Shatz (Eds.), *Blackwell handbook of language development* (pp. 87-106). Malden, MA: Blackwell Publishing.

Baron-Cohen, S., Baldwin, D. A., & Crowson, M. (1997). Do children with autism use the speaker's direction of gaze strategy to crack the code of language? *Child Development*, *68*, 48-57.

Bates, D., Maechler, M., & Bolker, B. (2012). lme4: Linear mixed-effects models using S4 classes [Computer software]. Available from: http://CRAN.R-project.org/package=lme4.

Berko, J. (1958). The child's learning of English morphology. *Word*, *14*, 150-177.

Berman, R. (2009). Beyond the sentence: Language development in narrative contexts. In E. Bavin, (Ed.) *Handbook of Child Language* (pp.354-375). Cambridge University Press: Cambridge.

Bloom, L. (1998). Language acquisition in its developmental context. In W. Damon (Ed.) *Handbook of child psychology: Vol. 2: Cognition, perception, and language* (pp. 309-370). John Wiley & Sons: Hoboken, NJ.

Boucher, J. (2012). Structural language in autistic spectrum disorder – characteristics and causes. *Journal of Child Psychology & Psychiatry, 53*, 219-233.

Brown, R. (1973). *A first language: The early stages*. Cambridge, MA: Harvard University Press.

Cameron, A. C. & Trivedi, P. K. (1998). Regression analysis of count data. *Econometrics Society Monograph*, 30. Cambridge University Press.

Carlson, M. T., Demir, E., Goldin-Meadow, S., Levine, S. (2013, April). *Using the path of MLU to predict later syntactic development*. Poster presented at the biennial meeting for the Society for Research in Child Development, Seattle, WA.

Carpenter, M., Nagell, K., Tomasello, M., Butterworth, G., & Moore, C. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the society for research in child development*. 63(4, No. 255).

Carrow-Woolfolk, E. (1985). *Test of auditory comprehension of language-third edition (TACL)*. Austin, TX: Pro-Ed.

Charman, T., Baron-Cohen, S., Swettenham, J., Baird, G., Drew, A., Cox, A. (2007). Predicting language outcome in infants with autism and pervasive developmental disorder. *International Journal of Language and Communication Disorders, 38*, 265-285.

Charney, R. (1980). Speech roles and the development of personal pronouns. *Journal of Child Language, 7,* 509-528.

Chawarska, K. & Shic, F. (2009). Looking but not seeing: Atypical visual scanning and recognition of faces in 2 and 4-year-old children with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 39*, 1663-1672.

Cheng, M., Khetrapal, N., Demuth, K., Fein, D. A., & Naigles, L. (2012). *Longitudinal changes in pronoun reversal in children with ASD and TD children.* Poster presented at the annual International Meeting for Autism Research, Toronto, ON.

Chomsky, N. (1959). A review of B. F. Skinner's "Verbal behavior". *Language, 35*, 26-58.

Chomsky, N. (1981*). Lectures on government and binding*. Dordrecht: Foris.

Cooper, R. P., & Aslin, R. N. (1990). Preference for infant-directed speech in the first month after birth. *Child Development, 61*, 1584-1595.

Crain, S., & Lillo-Martin, D. C. (1999). *An introduction to linguistic theory and language acquisition.* Oxford: Blackwell.

Cruttenden, A. (1977). The acquisition of personal pronouns and language'simplification'. *Language and Speech*, *20*, 191-197.

Elliot, C. (2007). *Differential ability scales – II*. Pearson Assessments: San Antonio.

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

Eigsti, I.-M., & Mayo, J. (2011). Implicit learning in ASD. In D. Fein (Ed.) *The neuropsychology of autism* (pp. 267-279). Oxford University: New York.

Evans, K. E. & Demuth, K. (2012). Individual differences in pronoun reversal: Evidence from two longitudinal case studies. *Journal of Child Language, 39*, 162-191.

Fenson, L., Dale, P., Reznick, J., Bates, E., Thal, D., & Pethick, S. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, 59.

Fombonne, E. (2005). The changing epidemiology of autism. *Journal of Applied Research in Intellectual Disabilities, 18*, 281-294.

Gentner, D. 1982. Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In S. Kuczaj (Ed.) *Language Development, Vol. 2: Language, thought, and culture* (pp. 301-334). Erlbaum: Hillsdale, NJ.

Gentner, D. & Boroditsky, L. (2001). Individuation, relativity, and early word learning. In M. Bowerman & S. Levinson (Eds.) *Language Acquisition and Conceptual Development* (pp. 215-256). Cambridge University Press: Cambridge.

Gerenser, J. (2009). Language disorders in children with autism. In R. G. Schwartz (Ed.) *Handbook of Child Language Disorders* (pp. 67-89). Psychology Press: New York.

Gernsbacher, M. A., Stevenson, J. L., Khandakar, S., & Goldsmith, H. H. (2008). Why does joint attention look atypical in autism? *Child Development Perspectives, 2,* 38-45.

Gibson, W. (1957). *The miracle worker: A play in three acts*. Knopf: New York.

Gillette, J., Gleitman, H., Gleitman, L., & Lederer, A. (1999). Human simulations of vocabulary learning. *Cognition, 73,* 135-176.

Gleitman, L. R. (1990). The structural sources of verb meanings. *Language Acquisition, 1*, 3–55.

Gleitman, L. R., Cassidy, K., Nappa, R., Papafragou, A., & Trueswell, J. C. (2005). Hard words. *Language Learning & Development, 1*, 23–64.

Golinkoff, R. M., Mervis, C. B., & Hirsh-Pasek, K. (1994). Early object labels: The case for a developmental lexical principles framework. *Journal of Child Language*, *21*, 125-125.

Gomez, R. L., & Gerken, L. (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition, 70*, 109-135.

Goodwin, A., Fein, D., & Naigles, L. R. (2012). Comprehension of wh-questions precedes their production in typical development and autism spectrum disorders. *Autism Research, 5*, 109-123.

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

Hirsh-Pasek, K., & Golinkoff, R. M. (1996). *The origins of grammar: Evidence from language comprehension*. Cambridge: MIT Press.

Hoff, E., & Naigles, L. (2002). How children use input to acquire a lexicon. *Child Development*, *73,* 418-433.

Hoff, E. (2006). How social contexts support and shape language development. *Developmental Review*, *26,* 55-88.

Hollich, G. J., Hirsh-Pasek, K., & Golinkoff, R. M. (2000). Breaking the language barrier: An emergentist coalition model for the origins of word learning. *Monographs of the Society for Research in Child Development*, *65,* (3, Serial No. 262).

Howard, A. A., Mayeux, L., Naigles, L. R. (2008). Conversational correlates of children's acquisition of mental verbs and a theory of mind. *First Language, 28,* 375-402.

Jordan, R. R. (1989). An experimental comparison of the understanding and use of speaker-addressee personal pronouns in autistic children. *International Journal of Language & Communication Disorders*, *24,* 169-179.

Karmiloff-Smith, A. (1998). Development itself is the key to understanding developmental disorders. *Trends in Cognitive Sciences, 2*, 389-398.

Kjelgaard, M. M., & Tager-Flusberg, H. (2001). An investigation of language impairment in autism: Implications for genetic subgroups. *Language and Cognitive Processes*, *16*, 287-308.

Landau, B., Smith, L. B., & Jones, S. S. (1988). The importance of shape in early lexical learning. *Cognitive Development*, *3*, 299-321.

Lee, A., Hobson, R. P., & Chiat, S. (1994). I, you, me, and autism: An experimental study. *Journal of Autism and Developmental Disorders*, *24*, 155-176.

Levy, S. E., Mandell, D. S., & Schultz, R. T. (2009). Autism. *Lancet*, *378*, 1627-1628.

Lidz, J. (2007). The abstract nature of syntactic representations: Consequences for a theory of learning. In E. Hoff & M. Shatz (Eds.), *Blackwell handbook of language development* (pp. 277–303). Malden, MA: Blackwell.

Liszkowski, U., Carpenter, M., Henning, A., Striano, T., & Tomasello, M. (2004). Twelve-month-olds point to share attention and interest. *Developmental Science, 7*, 297-307.

Lord, C., Risi, S., Lambrecht, L., Cook, E., Leventhal, B., DiLavore, P. C., Pickles, A., Rutter, M. (2000). The Autism diagnostic observation schedule-generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders*, *30*, 205-223.

Loveland, K. A. & Landry, S. H. (1986). Joint attention and language in autism and developmental language delay. *Journal of autism and developmental disorders*, *16*, 335-349.

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk*. *3rd ed*. Mahwah, NJ: Lawrence Erlbaum Associates.

Marcus, G. F. (1993). Negative evidence in language acquisition. *Cognition, 46*, 53–85.

Mayo, J., & Eigsti, I.-M. (2010). Brief report: A comparison of statistical learning in school-aged children with high functioning autism and typically developing peers. *Journal of autism and developmental disorders*, *42*, 2476-2485.

Moore, C. E., & Dunham, P. J. (1995). *Joint attention: Its origins and role in development*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Mullen, E. (1994). *The Mullen scales of infant development*. Circle Pines, MN: American Guidance Service.

Mundy, P. (1995). Joint attention and social-emotional approach behavior in children with autism. *Development and Psychopathology*, *7*, 63-63.

Mundy, P., Block, J., Delgado, C., Pomares, Y., Van Hecke, A. V., & Parlade, M. V. (2007). Individual differences and the development of joint attention in infancy. *Child Development*, *78,* 938-954.

Naigles, L. R. (1990). Children use syntax to learn verb meanings. *Journal of Child Language, 17*, 357-374.

Naigles, L. R. & Bavin, E. L. (2013). Introduction: Special issue on atypical development. *Journal of Child Language*, *40,* 1-10.

Naigles, L. R., Cheng, M., Khetrapal, N., Rattanasone, N. X., Fein, D., & Demuth, K. (submitted). "You're telling me!" The prevalence and predictors of pronoun reversals in children with autism spectrum disorders and typical development.

Naigles, L. R., Hoff, E., & Vear, D. (2009). Flexibility in early verb use: Evidence from a multiple-n diary study. *Monographs of the Society for Research in Child Development, 74*(2, No. 293).

Naigles, L. R., Kelty, E., Jaffery, R., & Fein, D. (2011). Abstractness and continuity in the syntactic development of young children with autism. *Autism Research, 4,* 1–16.

Naigles, L. R. & Swensen, L. D. (2007). Syntactic supports for word learning. In: E. Hoff & M. Shatz (Eds.) *The handbook of language development* (pp. 212-231). New York: Blackwell.

Naigles, L. R., & Tovar, A. T. (2012). Portable intermodal preferential looking (IPL): Investigating language comprehension in typically developing toddlers and young children with autism. *Journal of Visualized Experiments: JoVE*, *70*.

Park, C. J., Yelland, G. W., Taffe, J. R., & Gray, K. M. (2012). Morphological and syntactic skills in language samples of preschool aged children with autism: Atypical development? *International Journal of Speech-Language Pathology, 14,* 95-108.

Park, J. (2013). Early joint attention predicts children's later performance on comprehension tasks. (Master's thesis). Retrieved from Digital Commons at UConn.

Park, J., Tek, S., Fein, D. A., & Naigles, L. R. (2012, May). *Early joint attention predicts children with ASD's subsequent performance on comprehension tasks*. Poster presented at the annual International Meeting for Autism Research, Toronto, ON.

Paul, R., Chawarska, K., Cicchetti, D., & Volkmar, F. (2008). Language outcomes of toddlers with autism spectrum disorders: a two year follow-up. *Autism Research, 1*, 97-107.

Perner, J., Leekam, S. R., & Wimmer, H. (1987). Three-year-olds' difficulty with false belief: The case for a conceptual deficit. *British Journal of Developmental Psychology*, *5*, 125-137.

Pinker, S. (1994). *The language instinct*. New York: Harper.

R Studio. (2012). RStudio: Integrated development environment for R (Version 0.97.336) [Computer Software]. Boston, MA.

R Core Team. (2013). R: A language and environment for statistical computing. [Computer software.] R Foundation for Statistical Computing, Vienna, Austria. Available from: http://www.R-project.org/

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

Rollins, P. R. & Snow, C. E. (1998). Shared attention and grammatical development in typical children and children with autism. *Journal of Child Language*, *25,* 653-673.

Rowe, M. L., Raudenbush, S. W., & Goldin-Meadow, S. (2012). The pace of vocabulary growth helps predict later vocabulary skill. *Child Development, 83*, 508-525.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274,* 1926-1928.

Schopler, E., Reichler, R.J., & Renner, B.R. (1988). *The childhood autism rating scale*. Los Angeles, CA: Western Psychological Services.

Shulman, C., & Guberman, A. (2007). Acquisition of verb meaning through syntactic cues: A comparison of children with autism, children with specific language impairment (SLI) and children with typical language development (TLD). *Journal of Child Language, 34*, 411-423.

Siller, M., & Sigman, M. (2008). Modeling longitudinal change in the language abilities of children with autism: parent behaviors and child characteristics as predictors of change. *Developmental Psychology, 44,* 1691–704.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence.* Oxford University Press: USA.

Smith, L. B. (2000). Learning how to learn words. In R. M. Golinkoff, K. Hirsh-Pasek, L. Bloom, L. B. Smith, A. Woodward, N. Akhtar, M. Tomasello, & G. Hollich, (Eds.), *Becoming a word learner: A debate on lexical acquisition* (pp. 51-80). New York: Oxford University Press.

Snedeker, J., & Gleitman, L. R. (2004). Why is it hard to label our concepts. In D. G. Hall & S. R. Waxman (Eds.) *Weaving a lexicon* (pp. 257-294). Cambridge, MA: MIT Press

Snyder, W. (2007). *Child language: The parametric approach.* Oxford, UK: Oxford University Press.

Sparrow, S. S., Balla, D. A., & Cicchetti, D. V. (1984). Vineland adaptive behavior scales. Circle Pines, MN: American Guidance Service.

Stone, W. L., Coonrod, E. E., & Ousley, O. Y. (2000). Brief report: Screening tool for autism in two-year-olds (STAT): Development and preliminary data. *Journal of Autism and Developmental Disorders, 30*, 607-612.

Swensen, L. D., Kelley, E., Fein, D., & Naigles, L. R. (2007). Processes of language acquisition in children with autism: Evidence from preferential looking. *Child development*, *78*, 542-557.

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

Tager-Flusberg, H. (1994). Dissociations in form and function in the acquisition of language by autistic children. In: H. Tager-Flusberg (Ed.) *Constraints on language acquisition: Studies of atypical children* (pp. 175-194). Hillsdale, NJ: Erlbaum Publishers.

Tek, S. (2010). *A longitudinal analysis of joint attention and language development in young children with autism spectrum disorders*. (Doctoral dissertation). UConn: Storrs, CT.

Tek, S., Jaffery, G., Fein, D., & Naigles, L. R. (2008). Do children with autism spectrum disorders show a shape bias in word learning? *Autism Research, 1*, 208-222.

Tek, S., Mesite, L., Fein, D., Naigles, L. (2013). Longitudinal analyses of expressive language development reveal two distinct language profiles among young children with autism spectrum disorders. *Journal of Autism and Developmental Disorders,* doi 10.1007/s10803-013-1853-4.

Tenenbaum, E. J., Shah, R. J., Sobel, D. M., Malle, B. F., Morgan, J. L. (2012). Increased focus on the mouth among infants in the first year of life: A longitudinal eye-tracking study. *Infancy, 18*, 534-553.

Theissen, E. D. (2009). Statistical learning. In E. Bavin (Ed.) *Cambridge Handbook of Child Language* (pp.35-50). Cambridge: Cambridge University Press.

Thomas, M. S., & Karmiloff-Smith, A. (2003). Modeling language acquisition in atypical phenotypes. *Psychological Review*, *110*, 762-763.

Tomasello, M. (1992a). The social bases of language acquisition. *Social Development, 1,* 67-87.

Tomasello, M. (1992b). *First verbs: A case study of early grammatical development*. Cambridge: Cambridge University Press.

Tomasello, M. (1999). *The cultural origins of human cognition*. Cambridge: Harvard University Press.

Tomasello, M. (2000). Do young children have adult syntactic competence? *Cognition, 74,* 209-253.

Tomasello, M. (2006). Acquiring linguistic constructions. In D. Kuhn & R. Siegler (Eds.), *Handbook of Child Psychology*. New York: Wiley.

Tomasello, M. (2009). The usage-based theory of language acquisition. In E. Bavin (Ed.) *Cambridge Handbook of Child Language.* Cambridge: Cambridge University Press.

Treffert, D. A. (2009). The savant syndrome: An extraordinary condition. A synopsis: Past, present, future. *Philosophical Transactions of the Royal Society B, 364,* 1351-1357.

Waxman, S. R. (2004). Everything had a name, and each name gave birth to a new thought: Links between early word learning and conceptual development. In D. G. Hall & S. R. Waxman (Eds.) *Weaving a lexicon* (pp. 295-335). Cambridge, MA: MIT Press.

Wimmer, H. & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition, 13,* 103-128.

Yuan, S., & Fisher, C. (2009). Really? She blicked the baby? Two-year-olds learn combinatorial facts about verbs. *Psychological Science, 20,* 619–626.