ORIGINAL PAPER

Residual Language Deficits in Optimal Outcome Children with a History of Autism

Elizabeth Kelley · Jennifer J. Paul · Deborah Fein · Letitia R. Naigles

Abstract This study examined whether language deficits persist even in children with optimal outcomes. We examined a group of children with prior diagnoses on the autism spectrum who had IQs in the normal range, were in age-appropriate mainstream classes, and had improved to such an extent that they were considered to be functioning at the level of their typically developing peers. Fourteen such children between the ages of five and nine were matched on age and sex with typically developing children, and were given a battery of 10 language tests to investigate their language abilities. Results indicated that while these children's grammatical capabilities are mostly indistinguishable from their peers, they are still experiencing difficulties in pragmatic and semantic language.

Keywords Outcome · HFA · Language · Semantics · Pragmatics

Poor language and communication skills comprise a defining feature of Autism Spectrum Disorder (ASD; APA, 2000), and most children with ASD receive interventions that focus largely on building these skills. Recently, the prognosis for children with ASD receiving intensive early

intervention appears to have improved substantially. It has been reported that such children show significantly reduced problematic behavior, and improved language, cognition, and social interaction skills (Harris & Handleman, 2000; Jocelyn, Casiro, Beattie, Bow, & Kneisz, 1998; Lord, 1996). However, current outcome research on these children has reported only global cognitive measures or classroom placement; no one has yet investigated the language outcomes of these children in detail. Such an investigation seems warranted because global measures can mask both abilities and deficits in specific areas; for example, it is well known that language difficulties remain in children with Down syndrome, particularly in the area of grammar, despite years of intervention (e.g., Fowler, 1990; Fowler, Gelman, & Gleitman, 1994). Moreover, the contributing factors to good versus poor outcomes cannot be firmly established until their links with specific aspects of stable language behavior have been demonstrated. In this paper, we report the first such investigation of the language-related developmental outcomes of a group of ''optimal-outcome'' children who had been diagnosed with ASD during their preschool years; they are tested in grade school, after the cessation of their intervention.

Research over the past 20 years has revealed that a number of factors contribute to the overall developmental outcomes of children with ASD (Fenske, Zalenski, Krantz, & McClannahan, 1985; Gabriels, Hill, Pierce, Rogers, & Wehner, 2001; Lovaas & Buch, 1997). Most of the relevant studies compare the same children at two different times, which are anywhere from 2.5 years to 7 years apart; thus, the children are typically 3–6 years of age at time one (close to the initial diagnosis and onset of treatment) and 5–13 years of age at time two (i.e., the outcome). Children entering treatment at younger ages have been found to have higher global IQs and/or better school placements at time

E. Kelley · J. J. Paul · D. Fein · L. R. Naigles
Department of Psychology, University of Connecticut, 406 Babbidge Road, Unit 1020, Storrs, CT 06269-1020, USA

E. Kelley (✉)
Department of Psychology, Queen's University, 62 Arch St, H351, Kingston, ON K7L 3N6, Canada
e-mail: kelleyb@post.queensu.ca

Present Address:
J. J. Paul
Department of Psychiatry, University of Colorado Health Sciences Center, Denver, CO, USA

two than children entering treatment at older ages (Bibby, Eikeseth, Martin, Mudford, & Reeves, 2002; Fenske, et al., 1985; Harris & Handleman, 2000). An advantage for time two IQ has also been found for children whose time one IQs were relatively higher (Gabriels, et al., 2001), and an advantage for time two standardized language scores has been found for children whose initial verbal abilities were stronger (Szatmari et al., 2000). Finally, children who received intensive behavioral treatment for longer hours and/or more months had higher IQs and better school placements than children who received treatment for fewer hours/months (Gabriels et al., 2001; Lovaas, 1987; McEachin, Smith, & Lovaas, 1993). Further scrutiny of these studies reveals that the variability of outcomes in children with ASD encompasses a huge range: some children demonstrate no change in their already low IQ scores from time one to time two whereas others have been found at time two to score within the normal range, and/or have improved to such a great extent that they can be mainstreamed into chronologically age-appropriate classrooms in elementary school (Fenske et al., 1985; Gabriels et al., 2001; Harris & Handleman, 2000; Lovaas, 1987; McEachin et al., 1993). Full understanding of the causes of this variability, though, is hampered by the fact that the measures used to evaluate ASD outcome have been global rather than specific; most have just assessed children's IQ and/or their school placement. Detailed examination of cognitive functioning at outcome, especially of language functions, is crucial to understanding the degree to which their cognitive impairments have truly resolved, and to providing sensitive measures that can be used in future studies on predictive factors.

The problem of assessing and accounting for language-related developmental outcomes is complicated by findings that different sub-areas of language are differentially affected in ASD. In a recent literature review, Tager-Flusberg (2001) identified three components of language, computational (i.e., grammar), semantic, and pragmatic, and concluded that only the last was consistently impaired in individuals with autism.[1] Thus, any assessment of language outcomes in ASD will need to distinguish the pragmatic aspects of language. Moreover, even within the grammatical and semantic components, specific strengths and weaknesses are evident. For example, whereas Tager-Flusberg et al. (1990) and Tager-Flusberg (1994) have found that children with ASD displayed patterns of Mean Length of Utterance (MLU) increase, pronoun case development, and question form development similar to

typically developmentally mental-age-matched controls (see also Waterhouse & Fein, 1982), children with ASD have also been found to use a much more limited range of morphological and syntactic forms in their spontaneous speech: They produce fewer prepositions, conjunctions, articles, verb tenses, and auxiliaries, and they use very little complex syntax, including embedded sentences, sentence complements, and relative clauses (Bartolucci, Pierce, & Streiner, 1980; Cantwell, Baker, & Rutter, 1978; Fein et al., 1996; Menyuk & Quill, 1985; Scarborough, Rescorla, Tager-Flusberg, Fowler, & Sudhalter, 1991). Within the lexical realm, children with ASD have exhibited the same type of semantic-prototype organization in categorization at the basic and superordinate levels as language-matched controls, and consistently perform relatively well on standardized vocabulary tests (Kjelgaard & Tager-Flusberg, 2001; Tager-Flusberg, 1985, 2001; Waterhouse & Fein, 1982). However, children with ASD, unlike typically developing controls, do not recall lists of words in meaning-based chunks, nor do they turn them into a sentence to make recall easier (Hermelin & O'Connor, 1970; Tager-Flusberg, 1988; see also Dunn, Gomes, & Sebastian, 1996). Furthermore, their verbal memory deficits are increasingly apparent with material of increasing semantic organization (e.g., digits to sentences to stories) (Fein et al., 1996) Thus, it is possible that children with autism understand and/or store the meanings of words differently from typically developing children (Dunn & Rapin, 1997).

Within the pragmatic realm, where individuals with autism routinely experience their most severe difficulties, consistency across studies is greater (Dunn & Rapin, 1997; Fein et al., 1996; Kjelgaard & Tager-Flusberg, 2001; Prizant, 1996; Tager-Flusberg, 1996). Children on the autism spectrum have particular difficulty in responding to questions, sharing information, and requesting information (Tager-Flusberg, 1996; Wetherby, Prizant, & Schuler, 2000). They have great difficulty with false belief tasks, which are designed to test their understanding of Theory of Mind (Baron-Cohen, Leslie, & Frith, 1985; Happe, 1995; Tager-Flusberg, 1996, 1999). They also have difficulty with the production of narratives, in that children with ASD have been consistently found to use more bizarre language in their narratives than do mental-age-matched controls (Loveland, McEvoy, Tunali, & Kelley, 1990; Losh & Capps, 2003), and frequently use fewer causal connectives and a narrower range of narrative devices (Capps, Losh & Thurber, 2000; Losh & Capps, 2003; Tager-Flusberg, 1996).

In sum, any study of the language-related developmental outcomes of children with ASD will need to assess multiple aspects of grammar, lexicon, and pragmatics in order to present a complete picture of the children's abilities and deficits. Moreover, a comparison of the research on ASD

---

[1] Prosody, and phonology in general, are also areas of difficulty for children with autism (Shriberg, Paul, McSweeney, Klin, & Cohen, 2001; Wolk & Giesen, 2000). Because our study did not address these areas, though, space constraints do not permit us to review them.

outcome and language reveals reciprocal gaps in these literatures: As previously mentioned, studies that focus on the developmental outcomes of children with ASD have not investigated any aspects of their language in detail. Studies that have investigated in detail the language of children with ASD have not studied children who have reached the ''optimal outcome'' status of being fully mainstreamed with no educational supports. A third set of studies, which focus narrowly on the effects of specific types of training, do discuss both type of intervention (e.g., discrete trial learning or milieu learning) and specific aspects of language (e.g., learning spatial prepositions); however, these studies typically only include a very few participants (*n*s ranging from 1 to 8) and only a narrow range of outcome span (8 days to 8 months) and outcome measures (see Goldstein, 2002, for a review).

The purpose of the current study is to begin to fill these gaps in the research literature. We investigate in detail the language outcomes of a group of children for whom we have both early diagnosis and treatment information. These children were previously diagnosed with ASD between one and five years of age, were treated in intensive behavioral programs for between one and four years, and are now considered to be functioning successfully in a typical school environment. We chose to focus on this ''optimal-outcome'' group for several reasons: First, children such as these have never been examined in detail for their language strengths and weaknesses. We wished to determine the degree to which their language appears as ''recovered'' as other aspects of their behavior. From a practical standpoint, it is important to discover whether such children might still need some specific remedial assistance. Second, choosing children who are functioning academically within the normal range minimizes the heterogeneity of the usual autistic sample. While these children cannot be considered to have been 'modal' children with autism, their increased homogeneity with regards to diagnosis, treatment, and general outcome can only benefit the comparison with typical children. Third, the increased age and functioning of these children enables us to assess more complex aspects of their grammar and lexicon. For example, we can investigate whether their grammatical development (e.g., mastery of multi-clause syntax) has kept pace with their vocabulary growth, and whether their vocabulary growth has also led to increased understanding of the power of words to categorize (e.g., categorical induction). These are abilities manifested by typical children of four to five years of age.

We tested these children's language with a battery of psycholinguistic and standardized tasks, which assessed morphological, syntactic, lexical semantic, and pragmatic functioning. The children were given standardized tests to enable comparisons of their abilities with our typical control group, as well as psycholinguistic tasks that had been specifically developed to tap underlying linguistic knowledge in preschool-aged typically developing children while minimizing performance demands. Some tasks examined those properties of language that have been disputed with this population, including morphology and complex syntax within the grammatical realm, and the induction of category properties based on shared labels within lexical semantics. Other tasks tapped areas of language in which deficits or strengths have been found consistently, such as mental verb discrimination (Kazak, Collis, & Lewis, 1997; Lord, 1996; Ziatas, Durkin, & Pratt, 1998), narrative production (Capps et al., 2000), verb argument structure (Tager-Flusberg et al., 1990) and Theory of Mind tasks. Although Theory of Mind tasks do not assess language ability directly, these tasks have been shown to be strongly related to language ability (Happe, 1995). Our control group was chosen to be typically developing children who were matched on age (12 of the 14 were also matched on receptive vocabulary), because the null hypothesis is that these children with a history of autism are indistinguishable from their typically developing peers (e.g., Lovaas, 1987; McEachin, et al., 1993). Each psycholinguistic task we have chosen elicits above chance performance with typically developing children aged four-to-five years; therefore, comparisons will be made with these findings from the literature as well.

Method

Participants

The children with a history of ASD (called here the ASD group) were recruited from the clinical files of the third author, a clinical and research neuropsychologist specializing in autism and related disorders. Extensive clinical files were searched to identify children with clear diagnoses of PDD-NOS, Asperger's, or Autistic Disorder who had achieved optimal outcomes. All of the children had originally been diagnosed on the PDD spectrum as toddlers and still met criteria for a PDD diagnosis when evaluated by the third author (see Table 1 and Appendix A).

Table 1 indicates the age at which the third author first evaluated the child and the diagnosis given, as well as their current age. The children were diagnosed according to DSM-IV criteria (using a checklist with all DSM-IV symptoms) using extensive parent interview, child testing, and child observation.

In the Appendix there is a description of the DSM-IV symptoms displayed at that time, and notes on language development. The description in the Appendix indicates whether formal criteria were met for Autistic Disorder or

Table 1 Diagnostic information from third author

Child (random initials)	Age at Dx by author in months	Diagnosis given by author at this time	Original diagnosis	Age at current testing in months
T.C.	38	PDD-NOS	Autistic disorder	60
C.F. (f)	13	PDD-NOS	PDD-NOS	65
D.K.	64	Asperger's	Autistic disorder	77
P.B. (f)	44	PDD-NOS	Autistic disorder	79
M.K.	49	PDD-NOS	Autistic disorder	80
C.B.	58	Autistic Disorder	PDD-NOS	82
N.E.	77	PDD-NOS	Autistic disorder	86
L.X.	52	PDD-NOS	PDD-NOS	91
M.S.	51	PDD-NOS	Autistic disorder	93
T.S.	54	PDD-NOS	PDD-NOS	95
T.M.	73	Asperger's	Asperger's	98
U.N.	83	PDD-NOS	PDD-NOS	100
T.B.	90	Autistic Disorder	Autistic disorder	107
Q.K.	92	PDD-NOS (mild)	PDD-NOS	109

f = female

See appendix for a full description of these children's autistic symptoms and language functioning at the time given here

not at the time of the third author's evaluation. Eight children met formal criteria for Autistic Disorder and six for PDD-NOS. Note that in several cases, formal criteria for Autistic Disorder were met but a diagnosis of PDD-NOS or Asperger's Disorder was given to the family. The reason for this was that the child had previously been given a diagnosis of PDD-NOS or Asperger's and had continued to improve so that symptoms were fewer or milder than they had been. It was considered to be counterproductive to give the parents a more severe diagnosis when the child had actually shown significant improvement, and in many cases, the symptoms were present but in relatively mild form. For the youngest child (13 mos.) a

diagnosis of PDD-NOS was given; symptoms were not mild, but some symptoms (conversation, stereotyped play, stereotyped language, resistance to change) were not developmentally appropriate to score.

All of the children underwent intensive intervention programs at an early age, consisting of Applied Behavior Analysis (ABA), except for the one child with Asperger's syndrome (see Table 2 for the children's age of initiation and termination of treatment, and type of treatment).

At the time of the study, the children had all been mainstreamed into chronological age-appropriate classrooms, and were considered to be functioning at the level of their peers by parents and teachers. In many cases, the

Table 2 Participant characteristics

Child	Age at initiation of treatment*	Age at termination of treatment	Type of Program (ABA unless otherwise noted) and hours per week
T.C.	2 years	4 years	10 at school
C.F.(f*)	1 year, 2 months	4 years, 5 months	40 at home/school
D.K.	5 years	Ongoing at home	10 at school, 10 at home
P.B.(f*)	3 years	6 years	30 at school
M.K.	4 years, 5 months	5 years, 6 months	20 at school
C. B.	2 years, 9 months	6 years	20 at school, 15 at home
N.E.	3 years	5 years, ongoing at home	40 at home/school, then a few hours at home
L.X.	3 years, 1 month	5 years, 1 month	40 at home/school
M.S.	3 years	6 years	40 at home/school
T.S.	3 years	7 years	40 for 3 years, 15 for one year
T.M.	4 years	7 years	Special class at preschool
U.N.	3 years, 3 months	6 years, 4 months	40 at home/school
T.B.	4 years, 6 months	6 years	30 at home/school
Q.K.	3 years	6 years	Floor time/PECS/ABA 20 at home

* f = female

Approximate hours of treatment (hours per week × number of weeks)

Range = 1000–6750 h

Mean = 4054.62 h

SD = 2130.16 h

child's current teacher was not aware of the history of autism/PDD. Two children (see Table 2) continued to receive educational supports in the form of several hours per week of ABA at home to reinforce academics and work on language pragmatics but none received educational supports or ABA in school. At the time of the current study, the children ranged in age from 5 years 6 months to 9 years 1 month ($M = 7;3$, SD = 14 months). Twelve were boys and two were girls, a sex ratio that is typical for children with this disorder. All of the children resided in suburban or rural areas of Massachusetts.

The ASD group was matched on age and sex with a group of typically developing children (TD) (12 boys, two girls, mean age = 7;4, SD = 15 months; range 5;10 to 9;1). Although it was not our prior intention, the two groups ended up being matched on standard scores of the TACL vocabulary. The TD group was recruited through a primary school in suburban/rural Connecticut; the children's parents volunteered at a school PTO meeting where the purposes of the study were explained. The TD children were given small toys for their participation. All children were matched within a six-month age range of their counterparts in the experimental group. There were no significant differences between the groups in age ($t(26) = -0.50$, $p = .96$). All of these children were in age-appropriate classrooms and none of them were receiving special educational services at school.

Language assessments

The children were given 10 different language tests, described below:

The Test for Auditory Comprehension of Language, Third Edition (TACL-3) (Carrow-Woolfolk, 1985). This test assessed three categories of language understanding: Vocabulary, Grammatical Morphemes, and Elaborated Phrases and Sentences. Children were shown a page with three pictures on it, given a linguistic stimulus, and asked to point to the matching picture.

The Expressive One-Word Picture Vocabulary Test (EOWPVT) (Gardner, 1990). This task requires the child to name a pictured stimulus, providing a standardized measure of expressive vocabulary.

The Stanford-Binet Memory for Sentences Subtest (Thorndike, Hagel, & Satler, 1986) assessed children's verbal memory by asking children to repeat increasingly complex sentences.

The Wug Test of Productive Morphology (Berko, 1958). The Wug Test used nonsense words and pictures to assess children's ability to generalize the basic rules of English inflectional morphology. The 22 stimuli included six tests of the plural, nine of the past tense, four of the possessive,

two of the present tense, and one of the present progressive. All of the children received the test stimuli in the same order. They were shown a picture of an unusual fictional creature and told, ''This is a Wug.'' They were then shown another picture of the same type of creature and told, ''Now there are two of them. There are two _____.'' Children had to complete the statement by saying ''Wugs'' in order to receive credit for a generalization. Berko (1958) found that typically developing children as young as four years were able to make such generalizations regularly (see also Kim, Marcus, Pinker, Hollander, & Coppola, 1994).

Understanding of Complex Syntax (deVilliers & Roeper, 1995). This task investigated the extent to which the children understood how the syntactic phenomenon of ''*wh*-movement'' is affected by the presence of different types of subordinate clauses (e.g., Chomsky, 1982). That is, the *wh*-question, ''When did the girl say she planted the pumpkin?'' has two possible readings (i.e., When did she say?/When did she plant?), whereas the *wh*-question ''How did the mother learn what to bake?'' has only one (How did she learn?). The presence of the medial *wh*-word in the latter sentence promotes the reading of the matrix or main verb over the subordinate one. Our test sentences contained 12 embedded sentence complements, either with (six sentences), or without (six sentences) medial *wh*-words.

Children were told that they would be read some stories and at the end of each story there would be a question. The stories were told in the same order to all of the children. The responses were coded as to whether the children had answered the first or second clause of the sentence. deVilliers and Roeper (1995) reported that three-year-olds responded significantly differently to questions with and without the medial *wh*-word.

Verb Argument Structure (Naigles, Gleitman, & Gleitman, 1993). This task assessed children's understanding of verb argument structure by investigating how they enacted ungrammatical sentences, that is, sentences in which there were too many or too few noun arguments. Children were given a wooden ark and a number of wooden animals that they used to act out sentences spoken by the experimenter. The children were given twenty test sentences to enact; twelve were grammatical and eight were ungrammatical. The grammatical sentences (e.g., ''The elephant pushes the zebra.'') ensured that children understood the task: those who did not act out at least eighty percent of the grammatical sentences correctly were eliminated from the analyses.[2] Of the eight ungrammatical sentences, four included transitive verbs in intransitive frames (e.g., *The lion brings*), and four included intransitive verbs in

---

[2] This occurred with only one of the ASD group. In addition, there were technical difficulties with the videotape for one of the TD children; thus his enactments are not included.

transitive frames (e.g., *The tiger comes the horse*). The sentences were read in the same order for all the children.

Coding was performed as previously reported (Naigles et al., 1993; Naigles, Fowler, & Helm, 1992). The children's enactments of the ungrammatical sentences were coded as either: (a) *Frame Compliant*, in which the children would enact the sentence according to the meaning of the frame (the immature response), (b) *Verb Compliant*, in which the children would enact the sentence according to the meaning of the verb (the mature response), or (c) *Other*, used when the child either seemed to mishear the experimenter (e.g., using a *combing* motion when they heard the verb *come*) or used the wrong animals, or when it was unclear what the child was doing.[3] Only seven (out of 104) enactments were coded as Other; this is comparable to other results for children in this age group (Naigles et al., 1992). Naigles et al. (1992, 1993) found that children aged two to four years were more likely to behave *Frame Compliantly* whereas children five years and older were increasingly likely to behave *Verb Compliantly*.

Categorical Induction (Gelman & Markman, 1986). This task asked children to make a prediction about the property of a depicted natural kind object. The major question concerned whether the prediction (i.e., induction) would be made on the basis of *name* similarity. Eight natural kinds were examined, four of which were inanimate (e.g., a rock) and four animate (e.g., a rabbit). For each natural kind, children were shown the picture, told its name, and then told one of its properties. Except for the first two children tested (C.B. & U.N.), the children were asked to repeat back the name of the object and its special property to ensure they were paying attention. Then they were asked whether the next items also had that property. Four test cards were presented for each natural kind: (1) the exact same natural kind with a slightly different perceptual appearance (coded as *same*); (2) one of the same natural kind but of a different color or form *(target)*; (3) a perceptually similar object which was not of the same kind *(perceptual)*; and (4) a distracter item which did not bear any relation to the original natural kind other than whether

it was animate or inanimate *(distracter)*. Children were required to answer yes or no for each test card; if they said ''maybe'' they were asked to take their best guess. When the child refused to commit to an answer ($n = 4$ instances) these were always scored as incorrect. Both the order within and across test trials was counterbalanced across participants. Gelman and Markman (1986) reported that three-year-old and adults consistently make predictions based on name similarity.

Certainty differences with mental state verbs (Moore, Bryant, & Furrow, 1989). This task ascertained children's understanding of the certainty differences between *think* and *guess* as opposed to *know*. The materials included two puppets, a cat and a cow, and a number of white, yellow, and blue boxes. A three-trial pretest ensured that the children could distinguish the utterances of the two puppets and use these utterances to select the location of each sticker. All of the children responded correctly in at least two out of the three pretest trials.

During the test trials the child was presented with two boxes on each trial and given clues by the puppets. The puppets used the verbs *think*, *guess*, and *know* to designate the likely location of the sticker. For example, the cat puppet said, '' I guess it's in the blue box,'' whereas the cow puppet said, ''I know it's in the white box.'' Children were again asked to select the box holding the sticker. They were then told to put their box choice off to the side and wait until the end of the 12 trials to discover how many stickers they had obtained. The children all received the test trials in the same order. Four tests trials contrasted *think* and *know*, four contrasted *think* and *guess*, and four contrasted *guess* and *know*. Only the *think/know* and *guess/know* trials were tabulated for this analysis, as Moore et al. found that the *think/guess* distinction was not reliably understood until the age of eight. Moore et al. (1989) reported that five-year-old performed well above chance on the *think/know* and *guess/know* distinctions.

Theory of Mind tasks (Wimmer & Perner, 1983; *Perner, Leekham, & Wimmer,* 1987). Theory of Mind tasks are designed to assess whether children understand that other people can have a false belief. Two different tasks were used to assess the children's Theory of Mind capabilities.

The first of these tasks was called the Unexpected Location task and was a variation of the classic Maxi task (Wimmer & Perner, 1983). Children were introduced to a puppet, Astro, who said that she was going on a trip and needed to take her toy monkey with her. Astro then asked the children to help by putting her toy monkey in a blue box, which would function as her suitcase. She then remarked that she had forgotten snacks for her trip and left for the store. The experimenter then conspiratorially asked the children if they would like to play a trick on Astro, and the children and the experimenter would then move the toy

---

[3] For the intransitive verbs in transitive frames, *Frame Compliant* enactments were those where the child treated the first animal spoken as a causal agent of the second animal's motion (e.g., ''*the zebra comes the horse''*, enacted as the child using the zebra to push, force, or carry the horse). *Verb Compliant* enactments were those where the children had the horse and zebra move separately. For the transitive verbs in intransitive frames, *Frame Compliant* enactments were those where the children had the animal performing some sort of non-causative action (e.g., ''*The elephant brings*,'' enacted as simply moving the elephant around in some manner without using another object). If children added a direct object, and had the elephant move another animal from one location to another, the enactment was then coded as *Verb Compliant*. Enactments were coded as *Other* when they did not fit into one of the two preceding categories.

monkey from the blue box to the white box. The children were then asked the target question of where Astro would look for the monkey when she came back from the store. The children were also asked two control questions: where the monkey was now, and whether or not Astro had seen the monkey being moved. In the original task (Wimmer & Perner, 1983), children simply viewed puppets engaging in deception and being deceived. The current version of the task allowed much more involvement of the children. It should be noted, however, that this version of the task nonetheless has been found to yield similar developmental results (Mayeux, 2000).

The second Theory of Mind task was called the Unexpected Contents task (Perner et al., 1987). Children were shown a typical metal ''band-aid'' box and asked what they thought was inside; every child in this study said ''band-aids''. The experimenter then opened the box and showed the children that there were really balloons inside. The children were given a balloon, the box was closed, and the children were asked what was really inside as a control question. The children were then asked two target questions: what they thought was inside before the box was opened, and if the box had been shown to their best friend, what their friend would have thought was inside the box. The children always received the unexpected location task before the unexpected contents task and all the test questions were given in the order noted for all participants. Both tasks were scored for overall scores, and also for only the target questions addressing their understanding of those and other's mental states.

Narrative Capability (Capps et al., 2000; Tager-Flusberg & Sullivan, 1995). This task asked the children to relate a wordless picture book. The wordless picture book used for this task was ''Frog, Where are You?'' (Mayer, 1969). This book contains a story about a boy who loses his pet frog and encounters many adventures on the search for his frog, which he finds at the end of the story. The child and the experimenter first looked through the book together, page-by-page, but silently. The experimenter turned the pages for the children at this point to ensure that they did not move through the pictures too quickly. Children were then asked to tell the story from the beginning, in their own words, looking again at the book. While children were telling the story the experimenter would nod his/her head in agreement or give non-descript positive feedback (e.g., ''mm-hmm'' or ''really?''). If the child did not give much information about one page, or hesitated before turning the page the experimenter would ask ''Anything else?'' No specific leading questions were asked of the children.

The stories related by the children were then transcribed from the videotapes and coded for a number of different variables. General lexical variables such as number of words (tokens), number of different words (types), and the type-token ratio were compared between groups. More grammatical aspects were also examined, such as the number of clauses, number of connectives, number of tense changes, as well as various grammatical omissions such as omitted determiners and pronouns. Finally, the stories were coded for pragmatic variables such as causal attributions, mental state verbs, and dramatic devices such as intensifiers and sound effects (see Table 3 for complete explanations of all Narrative variables examined).

Vineland Adaptive Behavior Scales (Sparrow, Balla, & Cicchetti, 1984). Parents of the ASD group were interviewed with the Vineland Adaptive Behavior Scales to ensure that the children were in fact functioning within the normal range in the Communication and Socialization domains.

Procedure

All of the children were tested in their homes over the course of two sessions. The two sessions for the ASD group occurred on the same day, at least three hours apart. The two sessions for the typically developing children (TD) occurred on two different days, ranging from 2 days to 31 days apart ($M = 10.29$ days, SD = 7.32). The average session lasted 52.48 min for the ASD group (SD = 9.98), and 45.81 min for the TD group (SD = 5.74) (this difference was not statistically significant). Four different orders of the tests were generated; participants were randomly assigned to one of these four orders.

The children were taken to a quiet room in their home, and sat at a table with the experimenter who administered the tests. Another experimenter was also present to score the standardized tests and manipulate the video camera. Parents were allowed to observe their child's sessions but were not permitted to participate. As they were accustomed, the children in the ASD group were reinforced liberally for staying on task in the testing sessions with various edible reinforcers. They were also given a token to put on a token board at the end of each test and were verbally praised in a non-directive manner as well (i.e., mild positive praise regardless of the correctness of their answers). The typically developing children were not deemed to need or be used to the same motivational enhancers; thus, only verbal praise was used with these children.

All of the tasks were recorded on videotape. With the exception of the TACL-3 (for which the experimenter noted the correctness of the answers as the tasks proceeded), all tasks were coded from the videotape. Ten percent of the data from each task was coded by two trained observers; reliability ranged from 92% to 100%, with an average of 98% reliability.

Table 3 Narrative variables

(1) Word Tokens: Number of overall words in the narrative.

(2) Word Types: Number of different words in the narrative.

(3) Type/Token Ratio: Types over Tokens-a measure of lexical fluency.

(4) Number of Clauses: Number of simple clauses. All clauses contained at least one verb.

(5) Connectives: Number of words used to connect one clause to another (e.g., *because, so, then, and then, finally,* etc.) Numbers 1–5 were combined to create a general lexical variable.

(6) Complex Connectives: Percent of connectives used other than ''then'' or ''and then''.

(7) Tense Changes: Number of times child changed tense while telling the story, for example, the child may have been telling the story in the past tense and switched to the present tense.

(8) Verb agreement errors: When the child used the wrong form of the verb, that is, one that did not agree with the number of the noun (e.g., *the boy think*)

(9) Over-regularizations: The child used a grammatical rule on an irregular form of a verb or an irregular pluralization (e.g., *the frog goed*)

(10) Anaphora: The child used a pronoun without it being clear whom the referent was referring to. For example, if the child was talking about the boy and the dog in one sentence, and then merely said *he* in the next sentence, it would be unclear whom the child was referring to (unless the child had established the dog was a girl of course).

(11) Omitted morphemes: this is a composite variable that examined the number of omitted: determiners, pronouns, prepositions, and auxiliaries. A sentence with an omitted determiner would be, ''*the boy thought dog was gone*'' where *dog* is not being used as a proper name so needs a determiner. A missing pronoun was counted when the child omitted either subject or object pronouns, e.g. , ''*and looked for his frog*'' where there should be a *he* in subject position. Omitted prepositions were coded for in instances such as, '' *And he went the forest*'', where the child should have inserted *into* or *toward* before *the forest*. Finally, an omitted auxiliary was coded for utterances such as ''*the boy going to find the frog*'', where the child needed to insert the auxiliary *is* or *was* before the verb. Numbers 6–11 were combined to create a grammatical errors composite variable.

(12) Intensifiers: Narrative devices used to stress particular aspects of the story, or make the story more dramatic. These include a) hedges (*I'm not sure the boy knew what he was doing*), negatives (*the boy did not know those were deer's antlers*), and repetition for dramatic effect (*the boy looked and looked and looked for his frog*).

(13) Emotion and mental state verbs: any of these spoken by the child was calculated.

(14) Sound effects: things such as shouting, and using onomatopoeia.

(15) Speech: Any time the child spoke for the boy, dog, or various other animals. Numbers 12–15 were combined to create a narrative devices composite variable

(16) Misinterpretation: instances where the child incorrectly understood the scene, such as thinking that the deer's antlers are branches of a tree.

(17) Repetition: repetition of information that is not done for dramatic effect.

(18) Unclear reference: When what the child is saying has nothing to do with what is on the page of the storybook, e.g., *boy, girl, dog, frog, boy, girl, dog, frog*. Numbers 16–18 were combined to create a misinterpretation of the story variable. The last two variables (numbers 19 and 20) were not included in any composite variables.

(19) Causal discussion: Any time the child discusses the cause of an action or occurrence, e.g., *the boy fell down from the tree because an owl flew out*

(20) Goal discussion: Any time the child discusses the goals of the characters, such as the overall goal of the boy looking for his frog

Results

The first data analyses examined whether the children in the TD and ASD groups performed differently on each task. Unpaired two-tailed *t*-tests were performed comparing the children in the two groups (see Johnson & Carey, 1998, for a similar strategy). Twelve of the children in each group could also be matched, within one point of their raw scores, on the receptive vocabulary subtest of the TACL-3. Paired-sample *t*-tests were also performed for each task on these vocabulary-matched individuals to examine the extent to which the ASD children performed more poorly than their TD language controls. The second set of analyses examined whether the ASD children's performance on each task was related to their Communication and Socialization subscale scores on the Vineland. As the scores on the psycholinguistic tasks are essentially raw scores, all correlations were performed on both the raw (controlling for age) and the standard scores of the standardized tests

and the Vineland. Correlations were performed to examine the relationship between estimated length of treatment and scores on the various tasks. In addition, correlations between all tasks performed by the children were analyzed separately by group to determine if the patterns of relations between the tasks would be similar between groups.

Standardized tests

Table 4 presents the standard scores of the standardized language tests for both groups.

The ASD group appeared to be performing similarly to their peers on the vocabulary portions of the standardized language tasks. Note, however, that on the TACL-3 Grammatical Morphemes and Elaborated Sentences subtests, and the Sentence Memory task, the ASD group performed more poorly than the TD group, although their standard scores were still within the normal range.

Table 4 Standardized IQ measures

Test standard scores	Mean (SD) ASD group $N = 13$	Mean (SD) TD group $N = 14$	Equality of variance	t-test probability
TACL-3 Vocabulary (mean = 10)	10.54 (3.26)	11.14 (2.63)	$F=.679\ p > .05$	$t(25)= -.533\ p > .05$
TACL-3 Grammatical Morphemes (mean = 10)	9.77 (3.61)	12.29 (1.90)	$F = 3.662\ p = .067$	$t(18) = -2.292{\sim}\ p < .05$
TACL-3 Elaborated Sentences (mean = 10)	9.00 (4.62)	12.57 (1.70)	$F = 8.228\ p < .01$	$t(15) = -2.707{\sim}\ p < .05$
Expressive Vocabulary (EOWPVT) (mean = 100)	108.08 (30.14)	118.14 (14.22)	$F = 2.754\ p > .05$	$t(25) = -1.123\ p > .05$
Stanford-Binet Sentence Memory (mean = 50)	47.79 (8.89)	54.86 (8.61)	$F = 0.022\ p > .05$	$T(26) = -2.137\ p < .05$

The Wug test of productive morphology

Both groups performed comparably on the Wug task. T-tests revealed no significant differences between the groups (ASD: $M = 61.69\%$ correct, SD = 24.36; TD: $M = 71.43\%$ correct, SD = 31.00). Note that in Berko's (1958) original study, the first graders obtained an average of 68.59% correct; thus, the current sample, with an average age of seven, are roughly comparable.

Complex syntax

Table 5a presents the percent correct on the Complex Syntax task across sentence types.

T-tests revealed no significant differences between the groups. Table 5a also shows the children's percent correct for the sentence complements involving medial *wh*-words and those without medial *wh*-words. Again, no significant differences were found between groups on either sentence type. Similarly to deVilliers and Roeper's (1995) findings, both groups had significantly more difficulty with those sentences including medial *wh*-words than those without (ASD $t(13) = -4.301$, $p < .001$; TD $t(13) = -3.970$, $p < .01$).

Children's answers were then analyzed to determine whether they addressed the question of the first or second phrase. These data are presented in Table 5b. Results demonstrated that when there was no *wh*-word to restrict interpretation of the question, both the ASD group and the TD group answered the second clause more of the time (i.e., when the girl *planted*, not when she *said* she planted). However, when the medial *wh*-word was present, *both* groups answered the first clause significantly more (i.e., when the dog *learned*, not what he *caught*) (ASD: $t(13) = -2.28$, $p < .05$; TD: $t(13) = -4.47$, $p < .001$).

Verb argument structure

Both groups of children were significantly more Frame Compliant with the transitive ungrammatical sentences than the intransitive ungrammatical sentences; this is a pattern that has consistently been seen across age groups (Naigles et al., 1992, 1993; Naigles, Fowler, & Helm, 1995; Naigles & Lehrer, 2002), and was significant for both groups (ASD: $t(12) = 4.43$, $p < .01$; TD: $t(12) = 9.90$, $p < .001$). The ASD group tended to be more Frame Compliant than the TD group with the ungrammatical intransitive sentences (ASD = 29.17% (35.09), TD = 8.3% (12.5); $t(20) = 1.90$, $p = .077$), and performed significantly more Frame Compliantly with the ungrammatical transitive sentences (ASD = 82.5% (16.87), TD = 66.67% (16.28); $t(20) = 2.23$, $p < .05$).

Table 5 (a) Percent correct on the complex syntax task. (b) percentage of children answering which clause

Percent Correct	Mean (SD) ASD group	Mean (SD) TD group	Equality of Variance	t-test
Percent correct on the complex syntax task				
Across all Sentences	57.74 (28.95)	72.62 (15.13)	$F = -3.254\ p > .05$	$t(26) = -1.704\ p > .05$
Sentences with medial *wh*-words	36.90 (37.08)	54.17 (29.01)	$F = 2.749\ p > .05$	$t(26) = -1.372\ p > .05$
Without medial *wh*-words	78.57 (30.96)	90.48 (12.60)	$F = 5.593\ p < .05$	$t(24) = -1.333{\sim}\ p > .05$
Percentage of Children answering which clause				
ASD group with no medial *wh*-words				
When did the girl say ___23%_____ she planted the pumpkin___77%____?				
with medial *wh*-words				
How did Rover learn____58% _____ what to catch___42%___?				
TD group with no medial *wh*-words				
When did the girl say ___29%_____ she planted the pumpkin__71%_____?				
with medial *wh*-words				
How did Rover learn____62%_____ what to catch___38%_____?				

~ corrected for unequal variance

Table 6 (a) Overall percent correct for variables in the categorical induction task. (b) Percent correct for animate items

	ASD Mean (SD)	TD Mean(SD)	Variance	t-test	Sig.
(a) Overall percent correct for variables in the categorical induction task					
Across items	74.04 (13.46)	82.03 (5.51)	$F = 8.602 \; p < .01$	$t(16) = -1.911\sim$	$p = .066$
Target	52.68 (28.66)	70.19 (15.76)	$F = 2.037 \; p > .05$	$t(25) = -1.945$	$p = .063$
Same	83.04 (19.98)	94.64 (6.42)	$F = 7.991 \; p < .01$	$t(16) = -2.069\sim$	$p = .055$
(b) Percent correct for animate items					
Overall	70.68 (19.64)	81.99 (12.91)	$F = 3.842 \; p > .05$	$t(26) = -1.801$	$p = .083$
Target	44.64 (38.20)	73.21 (22.92)	$F = 4.489 \; p < .05$	$t(21) = -2.400\sim$	$p < .05$
Same	82.14 (24.86)	96.43 (9.08)	$F = 13.50 \; p < .001$	$t(16) = -2.020\sim$	$p = .060$

Categorical induction

For the Categorical Induction task, the first analysis was conducted to compare the two groups' overall scores (see Table 6a). A *t*-test comparing the number of correct answers approached significance. As described above, for each natural kind, four objects were presented. Of greatest interest is whether or not the child is able to induce the properties of another object of the same natural kind (the *target* object). When the *target* objects alone were examined for differences between the groups, this comparison also approached significance. For the stimuli that were of the same natural kind and only slightly perceptually different (*same*), the *t*-statistic also approached significance, with the TD group obtaining more correct answers. The *perceptual* stimuli, which looked similar to the test stimuli but was of a different category, did not show significant differences between the groups, nor did the *distracter* stimuli.

Because of the reported differences in how autistic children view animate versus inanimate objects (Fay & Schuler, 1980), the data were collapsed into groups divided along this dimension, and then separated into *target* and *same* (*perceptual* and *distracter* items were not included because of the similarity of the ASD and TD responses in the overall analyses). *T*-tests between groups showed the same pattern as above with the animate stimuli: the ASD group performed significantly more poorly on the *animate target* items, and the *t*-test comparing answers on the *animate same* approached significance (Table 6b). There were no differences between the groups for any of the *inanimate* stimuli. Given this pattern, repeated measures *t*-tests were conducted to determine if the ASD group would show significantly more difficulty with animate as opposed to inanimate items; all were non-significant.

Mental state verbs

The ASD group performed significantly more poorly than the TD group on the Mental State Verb task ($t(26) = 3.02$, $p < .01$). The ASD group performed correctly on the questions an average of 58.04% of the time (SD = 26.68%) whereas the TD group performed correctly 86.61% of the time (SD = 23.24%).

Theory of mind

When examining all three of the answers combined for the Unexpected Location task (whether the puppet saw the deception, where the toy monkey is now, and where the puppet will look for the toy monkey when she comes back), there were no significant differences between the groups (see Table 7). When only the target question was examined, however, a different picture emerged; the typically developing children performed significantly better than the ASD group.

The Unexpected Contents task revealed similar results (see Table 7). The ASD group had a mean correct of 2.21 (out of three), with a standard deviation of 0.80, and a range from one to three. The TD group had a mean of 2.79 correct with a standard deviation of 0.58 and a range from one to three. These means were significantly different between the groups ($t(23) = -2.16 \; p < .05$), even after correcting for close to unequal variances ($F(2,26) = 3.81$, $p = .062$).

Table 7 Percent correct on theory of mind tasks

Test	ASD group Mean (SD)	TD group Mean (SD)	Equality of variance	t-test probability
Unexpected				
Location	85.67	95.33	$F = 4.858$	$t(20) = -1.155\sim$
Overall	(28.40)	(12.10)	$p < .05$	$p > .05$
Unexpected				
Location	71.00	100.00	$F = 57.78$	$t(13) = -2.280\sim$
Target	(46.90)	(0.00)	$p < .001$	$p < .05$
Unexpected				
Contents	73.67	93.00	$F = 3.811$	$t(26) = -2.162$
Overall	(26.73)	(19.30)	$p > .05$	$p < .05$
Unexpected				
Contents	73.67	93.00	$F = 3.811$	$t(26) = -2.162$
Target	(26.73)	(19.30)	$p > .05$	$p < .05$

~Corrected for unequal variances

Narrative

In accordance with past research (Capps et al., 2000; Loveland et al., 1990; Tager-Flusberg & Sullivan, 1995), three different types of variables were examined with regards to the children's narrative competence: general lexical variables, grammatical errors, and pragmatic variables. There were no significant differences on any of the general lexical variables such as number of words (tokens), number of different words (types), or the type-token ratio, nor were there any significant differences on any of the grammatical variables; the data from these measures can be requested from the authors. The general lexical variables and grammatical errors were analyzed both individually and as composite variables (see Table 3); neither individual nor composite analyses yielded any significant differences between groups.

Table 8 presents the findings by group for the pragmatic variables. No differences between the groups emerged on many types of narrative devices, including the number of intensifiers used, such as repeating for emphasis (e.g., ''he *searched* and *searched* and *searched* for his frog''), hedges (e.g., ''I *think* the boy was worried about his frog''), or negatives for dramatic effect (e.g., ''the boy *didn't know* that those were deer's antlers''). There were also no differences between the groups on the number of emotion and mental state verbs used in the narrative; both groups used very few. The ASD group demonstrated significantly more sound effects and attributed significantly more speech to the characters than the TD group; however, this was due to two of the children taking on the persona of the characters and 'barking' and 'ribbiting' and 'buzzing' throughout the task. When these two children were eliminated from the analyses examining sound effects and attribution of speech, these differences were no longer significant. These variables were combined to create a composite variable: number of narrative variables (see Table 3).

Group differences emerged most strongly with those pragmatic variables most relevant to telling the story. The ASD group was significantly less likely to give causal explanations for story happenings, and they were significantly less likely to discuss the goals and motivations of the characters. Moreover, the clarity of the reference of the children (i.e., what objects or actions they were discussing), and the number of times they needlessly repeated information yielded differences that approached significance. The ASD group was also significantly more likely to misinterpret the story and thus give incorrect information in their narratives (i.e., thinking that the dog had jumped out the window instead of falling out). These last three variables were also combined into a composite variable: difficulties in interpretation of story (see Table 3).

Paired sample *t*-tests on vocabulary-matched pairs

As was mentioned earlier, 12 of the ASD children were able to be matched within one point of their raw score on the TACL-Vocabulary subtest with children from the TD group. Paired-sample *t*-tests were then conducted on these twelve pairs to determine if the children would continue to differ on these tasks when matched on very basic language skills. The ASD group continued to perform more poorly than their vocabulary-matched peers on a number of tasks: the Unexpected Contents task ($t(11) = -2.569$, $p < .05$), the Mental State Verb task ($t(11) = -3.500$, $p < .01$), and the number of causal explanations given in the Narrative task ($t(11) = -3.130$, $p < .05$). In addition, a number of these paired sample tests approached significance, with the ASD group again performing more poorly than their peers: the Elaborated Sentences subtest of the TACL ($t(11) = -1.907$, $p = .08$), the False Belief target questions ($t(11) = -1.915$, $p = .08$), and the number of times the story was misinterpreted in the Narrative task ($t(11) = -1.980$, $p = .08$).

Between-task correlations

The results of the correlations between tasks for the separate groups can be seen in Tables 9 and 10. Both

Table 8 Narrative task-pragmatic variables

	ASD group Mean (SD)	TDC group Mean (SD)	Equality of variance	*t*-test probability
Number of intensifiers	1.00 (1.49)	0.77 (1.30)	$F = 0.353$ $p > .05$	$t(25) = 0.431$ $p > .05$
Emotion or mental state verbs	1.07 (1.39)	1.38 (1.90)	$F = 0.728$ $p > .05$	$t(25) = -0.493$ $p > .05$
Sound effects	3.43 (5.20)	0.31 (0.48)	$F = 15.03$ $p < .01$	$t(13) = 2.248\sim$ $p < .05$
Speech attributed to characters	3.71 (4.34)	0.62 (1.33)	$F = 19.170$ $p < .001$	$t(16) = 2.467\sim$ $p < .05$
Misinterpretation of story	2.43 (1.91)	1.00 (1.00)	$F = 2.829$ $p > .05$	$t(25) = 2.406$ $p < .05$
Repetition of previous info.	3.57 (5.29)	0.92 (1.55)	$F = 8.603$ $p < .01$	$t(14) = 1.736\sim$ $p = .09$
Unclear reference	6.29 (4.81)	4.23 (2.71)	$F = 2.869$ $p < .061$	$t(21) = 1.379\sim$ $p = .10$
Causal explanations	0.07 (0.27)	1.23 (1.30)	$F = 21.744$ $p < .001$	$t(13) = 3.152\sim$ $p < .01$
Discussion of goals	1.64 (1.60)	3.31 (1.97)	$F = 0.114$ $p > 05$	$t(25) = -2.417$ $p < .05$

~controlled for unequal variance

groups demonstrated a general pattern of correlations between tasks within aspects of language, for example, correlations between vocabulary tests and correlations between syntactic tasks. The main difference between the two groups was that the TD group showed correlations between the narrative tasks and other aspects of language competence (as well as within narrative tasks), whereas the ASD group showed within-narrative correlations only. That is, their narrative ability did not seem to be correlated with any other language tasks.

Vineland adaptive behavior scales

These scales were not given to the typically developing children. The children with a history of autism were completely within the normal to high range on the Communication subscale (Mean standard score = 98.92, SD = 13.18, range = 81–132) and all but two of the children scored within the low normal to normal range on the Socialization subscale (Mean standard score = 80.23, SD = 10.62, range = 63–101).

Correlations between the language tasks and the Vineland

Pairwise correlations were performed between the ASD children's standard scores on the Vineland Communication and Socialization subscales and their scores on each language task; only two reached significance: The Sentence Memory task ($r = .71$, $p < .01$) and the number of narrative devices used in the Narrative tasks ($r = .57$, $p < .05$) both correlated significantly with the Vineland Communication standard score. However, as mentioned earlier, two children in the ASD group took on the point of view of the characters and 'barked' and 'ribbited' throughout; when these two children were eliminated from the analysis, the latter correlation was no longer significant. When the *raw* scores of the Vineland Communication and Socialization subscales were used (controlling for age), the only significant correlation was between the Stanford-Binet Memory for Sentences raw score and the Vineland Communication subscale ($r = .79$, $p < .05$).

Correlations between the language tasks and estimated hours of treatment

Pair-wise correlations were performed between the children's scores on each language task discussed above and the number of hours of treatment they experienced. No significant correlations were obtained between language scores and the number of hours in treatment.

Correlations between the language tasks and number of early autistic symptoms

Correlations were also run between number of early autistic symptoms from the DSM-IV and the language tasks. The only significant correlations were between the False Belief task and the number of Social autistic symptoms ($r = -.60$, $p < .05$) and the number of Communicative symptoms ($r = -.56$, $p < .05$). The negative correlations indicate that the more autistic symptoms they had originally, the more poorly they did on the False Belief task.

Discussion

This study investigated the language abilities of grade-school-aged children who had a history of ASD but now functioned in the normal range on standardized measures of language and communication skills, and functioned without support in mainstream classes, and compared their abilities to their typically developing peers. The primary question was whether a comprehensive assessment of the grammatical, semantic, and pragmatic aspects of their language would reveal residual deficits in the ASD group. The results were threefold.

First, a solid area of strength for the ASD group was found within the standardized tests. These children did not differ from the TD group on the vocabulary assessments, and although they performed less well than the comparison group on the other subtests, they tested well within the normal range for their age. The significant differences, therefore, were probably due to the above-average performance of the TD group on these tests.

Second, the psycholinguistic tasks revealed both strengths and weaknesses in the ASD group's language. In particular, the ASD group's performance was indistinguishable from that of the TD group on the Productive Morphology and Complex Syntax tasks. However, the ASD group did perform somewhat more poorly than the TD group on the Verb Argument Structure task, and on the target questions, particularly when they involved animate stimuli, of the Categorical Induction task. The ASD group performed significantly more poorly than the TD group on the Mental Verb task, the Theory of Mind task, and the pragmatics aspects of the Narrative task.

Third, only one significant correlation emerged between the ASD children's scores on the Vineland and their performance on the language tasks; only their scores on the Sentence Memory measure correlated significantly with their scores on the Vineland Communication subscale. Children who had initially presented with more severe autistic features now performed more poorly on the False Belief task.

In sum, there were a number of language tasks on which the ASD group performed reliably more poorly than the TD group. Moreover, in each psycholinguistic task where the ASD group's performance suffered in comparison with the TD group, the ASD group's performance also seemed outside the normal range for their chronological (and indeed, receptive vocabulary) age. That is, the degree of Frame Compliance exhibited by the ASD group in the Argument Structure task was close to that demonstrated by typically developing preschoolers (i.e., over 70% (Naigles et al., 1993) rather than grade-schoolers (55–65% (Naigles et al., 1992)). Their percent correct on the target questions in the Categorical Induction task (52%) was much lower than that found for typically developing preschoolers (over 70%; Gelman & Markman, 1986). Their percent correct on the Mental Verb task was comparable to that of typically developing three-year-olds, as was their performance on

the Theory of Mind tasks and the Narrative variables (Berman, Slobin, Stromquist, & Verhoeven, 1994). Thus, the differences found between the ASD and TD groups are not wholly attributable to above-average or ceiling performance by the TD group; instead, the ASD group's performance on these tasks was more on a par with that of typically developing preschoolers. In sum, the results of this study indicate that these children with a history of ASD are not (yet) at a language level that is on a par with their chronological age-mates on specific functions.

What is the pattern of language abilities and deficits in these children with a history of ASD?

Despite the fact that the children with a history of autism told stories that were just as long and grammatically indistinguishable from those of their peers, they were still

Table 9 Correlations between tasks for ASD group

	TACL vocabulary standard score	TACL morphology standard score	TACL elaborated sentences standard score	EOWPVT standard score	Wug percent correct	Stanford-Binet sentence memory raw score
TACL morph standard score	.60*					
TACL elabor. sentences SS	.60*	.64*				
EOWPVT standard score	.78**	.37	.33			
Wug percent correct	.22	.42	.20	.55~		
S-B sentence Memory RS	.49~	.50~	.62*	.45	.45	
Categorical induction percent correct	.20	.67*	.32	.02	.29	.22
False belief target	.33	.34	.19	−.09	.10	.51~
Un. contents target	.40	.40	.22	.74**	.49~	.33
Mental state verb % cor	.51~	.58*	.38	.25	.24	.44
Complex syntax % cor.	.44	.68*	.66*	.18	.42	.61*
% Frame compliant transitives	−.61~	−.63~	−.25	−.68*	−.53	−.004
Percent frame compliant intransitives	.17	.48	−.11	.11	−.16	−.03
Narrative-general lexical var.	−.13	−.14	−.10	−.14	−.02	−.17
Narrative-grammatical errors	−.16	−.27	−.50	−.36	−.49~	−.45
Narrative-no. of devices	−.15	−.17	.18	−.28	−.002	.35
Narrative-difficulties	−.14	−.05	.24	−.19	−.11	.20
	Categorical induction percent correct	False belief target question correct	Unexpected contents percent of target questions	Mental state verbs percent correct	Complex syntax percent correct	Percent frame compliant on ungram. transitives
False belief target ques.	.37					
Unexpected contents	.33	−.03				
Mental verbs % correct	.61*	.43	.23			
Complex syntax	.39	..41	.23	.24		
Ungrammatic. transitives	−.55~	−.03	−.33	−.52	−.32	
Ungrammatic. intransitives	.10	.07	.07	−.25	.32	−.45
Narrative lexical	−.12	.13	−.15	−.04	.21	.23
Narrative-grammatical	−.17	.03	−.23	−.11	−.09	.29
Narrative-no. of devices	−.13	.27	−.38	.07	.09	.28
Narrative-difficulties	−.17	−.20	−.19	−.17	−.04	.32
	Frame com. intransitives	Narrative lexical	Narrative gram. errors	Narrative devices		
Narrative lexical	−.06					
Narrative grammatical errors	.16	.73**				
Narrative devices	−.39	−.38	−.46~			
Narrative difficulties	−.27	−.67**	−.57*	.74**		

~ p < .10; * p < .05; ** p < .01

experiencing difficulties in the pragmatic realm. For example, when retelling the Frog Story, our ASD group provided incorrect and redundant information, few causal explanations, and few mentions of the goals and motivations of the characters. Our ASD group also was more likely to interpret the Theory of Mind questions solely in terms of their own knowledge state rather than also taking into account the knowledge states of others. In both of these findings, we corroborate those already reported in the literature for ASD children (Fein, 2001; Geller, 1998; Happe, 1995; Lord & Paul, 1997; Tager-Flusberg, 1996, 1999): Thus, children with autism have consistent difficulty with the pragmatics component of language, and our children with a history of autism still manifested that difficulty.

The tasks more related to the grammatical sub-domain elicited generally better performance. These children demonstrated their ability to generalize morphological rules to novel items in the Productive Morphology task as well as their sensitivity to the syntactic ramifications of the presence of a medial *wh*-word in a multi-clause *wh*-question in the Complex Syntax task. The ASD group's good performance on this latter task contrasts with their poor performance on the Mental Verb task, which also involves multi-clause sentences. It seems more likely, though, that the Mental Verb task is tapping the children's pragmatic and/or semantic knowledge rather than their syntax, as the contrasting stimuli in this task had the same syntactic form but different semantic and pragmatic implications (i.e., that *guess* and *think* are less certain than *know*). Thus, it is more parsimonious to attribute the ASD group's poor performance on the Mental Verbs task to their already attested pragmatic (or semantic, see below) difficulties, rather than any syntactic difficulties.

The ASD group's performance with the Verb Argument Structure task also deserves mention. As described earlier, these children were significantly more Frame Compliant with the ungrammatical transitive sentences than the TD;

Table 10 Correlations between tasks for TD group

	TACL vocabulary standard score	TACL morphology standard score	TACL elaborated sentences standard score	EOWPVT standard score	Wug percent correct	Stanford-Binet sentence memory raw score
TACL morph standard score	.60*					
TACL elabor. sentences SS	.05	.50~				
EOWPVT standard score	−.68**	−.24	.09			
Wug percent correct	−.39	.01	.19	.54*		
S-B sentence memory RS	−.53*	.10	.29	.52~	.56*	
Categorical induction percent correct	−.05	−.24	.13	−.08	.21	.07
False belief target un. contents target	−.13	.34	.13	.33	.43	.29
mental state verb % cor	−.11	.22	.06	.13	.34	.58*
Complex syntax % cor.	−.46~	.09	.58	.31	.69**	.67**
% Frame compliant transitives	−.22	−.05	.06	.11	.20	.01
Percent frame compliant intransitives	.20	.17	−.21	−.13	.05	.02
Narrative-general lexical var.	−.28	.26	.13	.40	.77**	.34
narrative-grammatical errors	−.48~	−.03	−.05	.53~	.40	.50~
Narrative-no. of devices	−.30	.24	.10	.52~	.50~	.18
Narrative-difficulties	−.05	−.23	−.57*	−.21	−.05	−.42
	Categorical induction percent correct	False belief target question correct	Unexpected contents percent of target questions	Mental state verbs percent correct	Complex syntax percent correct	Percent frame compliant on ungram. transitives
False belief target ques. unexpected contents	−.47~					
Mental verbs % correct	.61*		.70**			
Complex syntax	.39		.30	.24		
Ungrammatic. transitives	−.55~		.000	−.14	.12	
Ungrammatic. intransitives	.10		.000	.50	−.38	.35
Narrative lexical	−.12		.44	.19	.46	−.04
Narrative-grammatical	−.17		.006	.07	.15	.14
Narrative-no. of devices	−.13		.24	−.05	.20	.24
Narrative-difficulties	−.17		.09	−.11	−.49~	−.20
	Frame com. intransitives	Narrative lexical	Narrative gram. errors	Narrative devices		
Narrative lexical	−.003					
Narrative grammatical errors	−.06	.57*				
Narrative devices	.34	.62*	.26			
Narrative difficulties	−.36	−.15	.05	−.24		

~ *p* < .10; * *p* < .05; ** *p* < .01

their scores resembled those of typically developing preschoolers (Naigles et al., 1993). It is likely, though, that this level of performance actually reflects *lexical* rather than syntactic immaturity. As discussed by Naigles et al. (1992, 1993, 1995,) this verb argument structure task taps syntactic knowledge insofar as it assesses whether participants understand the relation between sentence frames and sentence meanings; for example, that transitive frames canonically imply causative meanings (e.g., Jackendoff, 1990; Levin, 1993). Our finding that the ASD systematically interpreted the ungrammatical transitive frames causatively indicates that they are in possession of this grammatical knowledge. However, the task also taps lexical knowledge insofar as it assesses whether participants have mastered the verb-specific frame requirements of English; for example, that *bring* must be transitive and *go* must be intransitive. This, we conjecture, is the aspect of knowledge that our ASD group were having difficulty with. Previous findings have suggested that this aspect of verb argument structure acquisition is accomplished on a verb-by-verb basis in typically developing children (Naigles et al., 1992); it seems likely that the ASD group were further behind in this lexical process.

Finally, the ASD group also performed more poorly on the other task that tapped lexical semantics, namely, Categorical Induction. That is, these children had considerable difficulty in realizing that a given property of a named natural kind should ''transfer'' to another named instance of that natural kind (i.e., that if told a white rabbit eats grass, should induce that a brown rabbit eats grass, too). Thus, our findings are more similar to those of Dunn et al. (1996) and Hermelin and O'Connor (1970) than to those of Tager-Flusberg (1985). On the surface, these children's poor performance with the lexical semantics tasks is at odds with their excellent performance on the standardized vocabulary assessments. However, the children's performance on the standardized vocabulary tests indicated only that they *knew* many words, thus demonstrating understanding of the *identification function* of words and concepts. What they may have had difficulty with are other conceptual functions, such as that which motivates the relations between concepts and that which allows for the induction of hidden properties (e.g., Gelman & Medin, 1993).[4]

---

[4] Recall that, while the ASD group's scores on the Categorical Induction task were consistently lower than those of the TD, this difference only reached significance for those items that depicted animate natural kinds (see Table 4). Greater capability with inanimate than animate stimuli has been reported before in connection with ASD: Animate objects may be over-arousing; hence, less attention is paid to them (Frith & Happe, 1998) or they are less inherently reinforcing to children (Hauck, Fein, Maltby, Waterhouse, & Feinstein, 1998).

Taken together, those tasks that most closely tap children's pragmatic language (Theory of Mind, Narrative, Mental Verbs) elicited the poorest performance by our ASD group, whereas the grammatical tasks (Productive Morphology, Complex Syntax) yielded the most consistently high performance of any domain. Tasks that required children to use their lexical knowledge showed mixed findings with good overall vocabulary but impairments in specific aspects of lexical understanding. Thus, the psycholinguistic tasks do appear to cluster together to reveal a pattern of linguistic strengths and weaknesses in these children with a history of autism. The pattern replicates to a large extent that put forth by Tager-Flusberg (2001), who proposed that children with ASD would perform poorly within the pragmatics domain but quite well within the grammatical domain (but see Kjelgaard & Tager-Flusberg, 2001, for a discussion of a subgroup of children with autism with grammatical difficulties). However, whereas Tager-Flusberg (2001) considered the semantics abilities of individuals with autism to be fairly intact, we found that our ASD group showed mixed performance in the lexical semantics domain (see also Dunn & Rapin, 1997).

This conclusion is supported by the findings from the correlations performed between all of the tasks (Tables 9 and 10). Several patterns within the two correlation matrices are similar. For example, both groups yielded comparable numbers of correlations within the lexical domain (e.g., between TACL vocabulary and the EO-WPVT, between Categorical Induction and Mental Verbs), within the grammatical domain (e.g., between the Stanford-Binet sentence memory test and the test of Complex Syntax), and across the lexical and grammatical domains (e.g., between TACL morphology and Categorical Induction, between TACL vocabulary and the Stanford Binet sentence memory test). Moreover, both groups yielded few correlations between the language measures and the Theory of Mind-related measures (i.e., one each with the Unexpected Contents test, two each with Mental Verbs). Where the groups differed most involved the narrative measures: Whereas these were *inter*-correlated for each group, *only* the TD children's scores also yielded significant correlations *between* the narrative measures and other aspects of language[5] Thus, it appears that for the typically developing children, pragmatic ability, as manifested by their narratives, is well-integrated with other areas of language. In contrast, the pragmatic abilities of the ASD appear to stand alone.

---

[5] When the correlations were conducted using the raw scores of the standardized tests, the TD group's scores yielded three additional significant correlations between narrative variables and the lexical/grammatical measures; the ASD yielded none.

What predicts ASD children's performance on the language tasks?

There is little to suggest that the ASD children's adaptive skills were associated with their performance on the language tasks. Only one of the tasks yielded a significant correlation (once outliers were removed), the Sentence Memory task. The range of Vineland Communication standard scores (81–132) seems wide enough to support more correlations; thus, it is possible that the language tasks and the Vineland are simply tapping different abilities and deficits in the children. Although Vineland Communication measures adaptive use of communication, at the age and competence level of the children tested, it largely measures specific, taught skills, such as grade level for reading, using a dictionary, writing notes, and addressing envelopes and would not be expected to be sensitive to the kinds of deficits found here.

None of the correlations between the language tasks and the estimated amount of treatment were significant. Moreover, the children's early diagnostic symptoms only predicted their performance on the False Belief task. It is possible that the paucity of correlations is attributable to our small sample size. In order to examine the true effects of early severity and duration and type of treatment on cognitive outcomes, one would need a larger sample with greater variation in treatment, and early measures of cognition. It is also possible that children with more severe early symptoms received earlier and longer intervention, thus suppressing any correlations.

Conclusions

Several limitations to the current study should be mentioned. The sample size was relatively small and homogeneous; although it might be difficult to identify a large sample of equally successful children, such an effort would allow firmer conclusions about this group of children. The sample size was too small to allow examination of results by prior diagnosis, such as Autistic disorder vs. PDD-NOS, although examination of the Vineland scores suggests no differential communication outcome by early diagnosis. As mentioned above, treatment type and SES were also of limited variability. Second, the children received clinical but not research diagnoses (e.g., ADOS or ADI), although the clinical diagnosis was given by a diagnostician who has specialized in this area for many years (see Table 3 for the children's diagnostic markers). It should also be noted that the children's diagnoses were given long before the study was even conceptualized, thus making any bias

unlikely. Third, for the current study, the children were assessed for language and adaptive functioning; it was clear that most of them would no longer meet criteria for any form of PDD but formal diagnostic procedures were not administered at the time of testing. Fourth, no measure was given to examine the children's level of psychopathology. It is possible that these children are now manifesting a different disorder. Finally, no nonverbal tests were given to the children. Although our primary interest was the children's language status, a measure of nonverbal IQ might have been illuminating. A follow-up of these children, including formal diagnostic procedures, a parent-reported measure of psychopathology, and nonverbal IQ measures, is under way.

As mentioned in the introduction, this study is the first to investigate in detail the language abilities and residual weaknesses of children who were previously diagnosed with ASD and who have shown such great improvement that they are mainstreamed into age-appropriate classrooms in elementary school with minimal or no supports. On the one hand, the dramatic improvement of these children can be seen in the mere fact that they were able to participate fully in all tasks, split only over two hour-long sessions, and that their Vineland Communication and Socialization scores were all within the normal range. Moreover, their receptive and expressive vocabulary scores, which were not significantly different from those of the TD, attest to their strong ability to learn the mappings between words and their referents. Scores on the grammatical tests were either indistinguishable from their peers (i.e., Productive Morphology, Complex Syntax) or well within the normal range for their chronological age (Grammatical Morphemes, Elaborated Sentences, Verbal Memory). Taken together, these results paint a picture of children whose knowledge of grammar and vocabulary is appropriate for their age level.

However, their performance was not consistently at this level. Significant residual deficits in some areas of language still remain in these children, particularly with regard to lexical semantics and pragmatics. These results have clinical implications for children with a history of ASD. The language difficulties described above suggest that the children might benefit from periodic in-depth language assessments and continued language therapy where indicated. Standardized scores in the average range and adequate performance academically may have led some school systems to discontinue support services prematurely. The difficulty with complex lexical and pragmatic tasks suggests that social cognition may be an area of continuing difficulty. It is possible that as the children face the more complex social and linguistic demands of the higher grades, their academic and social adjustment may be at risk.

Appendix

T.C. was seen by the third author at 38 months. At this time he had mild nonverbal communication difficulties, problems in peer relations, a lack of emotional reciprocity, poor conversation skills, stereotyped language, stereotyped play, and a preoccupation with parts of objects. Format criteria for Autistic Disorder were not met because his nonverbal communication difficulties were too mild; he met the formal criteria for PDD-NOS. Regarding his language history, he had a few words at 18 months and two-word phrases at 20 months, but did not seem to comprehend much language and much of his expressive language was echolalic. He did not point until he was taught to do so at age two years. At 38 months he had a receptive language delay on the Preschool Language Scale of seven months, and his expressive language was nine months delayed. His Vineland Communication age equivalency at this time was two years, seven months; he was just beginning to relate experiences to others. T.C. was diagnosed with PDD-NOS on the basis of the following symptoms from the DSM-IV criteria: poor eye gaze, difficulty forming friendships, withdraws from social interaction, echolalia, stereotypies, and preoccupation with parts of objects.

C.F. was seen by the third author at 13 months. At this time she had nonverbal communication difficulties, no joint attention, a lack of emotional reciprocity, delayed language, and preoccupations. She met the formal criteria for PDD-NOS at this time. Regarding her language history, she had babbled from six to nine months and then suddenly stopped. At eleven months she showed a regression in eye contact, response to language, and social interaction. At thirteen months she made almost no vocalizations across the two hours of the testing session and demonstrated flat affect. At home she would vocalize only in whispers and jargons in monotone and she no longer responded to her name or other language although she would respond to gesture. Her testing at 13 months showed a Vineland Communication age equivalence of one to three months, a Socialization age equivalence of four to eight months, and an average Daily Living score. An SLP evaluation at thirteen months found interaction, receptive language, and expressive language to all be between the ages of three to six months, with a few higher scattered skills. C.F. was diagnosed with PDD-NOS on the basis of the following symptoms from the DSM-IV criteria: impaired eye gaze, lack of pointing and gestures, lack of social or emotional reciprocity, delayed language and repetitive behaviors.

D.K. was seen by the third author at 64 months. At this time he had nonverbal communication deficits, difficulties in peer relations, a lack of joint attention and emotional reciprocity, difficulties in conversation, stereotyped language and play, preoccupations and routines, and preoccupations with parts of objects. At this time he met the formal criteria for Autistic Disorder. Regarding his language history, he had spoken his first sentences at 23 months, but his language and play were both repetitive and scripted. In playgroup at ages two-and-a-half to three, he talked about other children but did not interact with them. When tested at 64 months, conversation was impeded by aggression, poor attention, and insensitivity to his conversational partner. His Preschool Language Scale showed scores in the superior range, and Vineland Communication scores were only slightly depressed (age equivalence was 5–0), but his Daily Living and Socialization scores were more than two years delayed. D.K. was diagnosed with Autistic Disorder on the basis of the following symptoms from the DSM-IV criteria: lack of gestures, poor eye gaze, failure to develop peer relationships, lack of pointing to share attention, lack of social emotional reciprocity, delayed echolalia, lack of make-believe play, delayed language, over-adherence to routines, preoccupation with parts of objects, and repetitive self-stimulatory behaviors.

P.B. was 46 months when seen by the third author. At this time she had deficits in nonverbal communication, problems in peer relations, stereotyped language and play, and routines. She met the formal diagnostic criteria for PDD-NOS at this time. Regarding her language history, she spoke her first single words at 10 months and two-word combinations at 12 months. At 30 months she had no conversation and would repeat or ignore others' questions. Symbolic play emerged at 42 months but was highly repetitive. At age three, an SLP evaluation noted atypical language and social skills, and poor social use of language. P.B. had pronoun reversal difficulties until 44 months. At 46 months P.B. did well on formal language testing except on items requiring responses to open-ended questions or word definitions. Her WPPSI Vocabulary scaled score was five, Comprehension was six, and Information was nine. Vineland Communication score was age equivalent (3–8), but Daily Living and Socialization were delayed by about a year. P.B. was diagnosed with PDD-NOS on the basis of the following symptoms from the DSM-IV criteria: impairment in eye-to-eye gaze, failure to develop peer

relationships, lack of pointing to express interest, poor social use of language, restricted play, and adherence to routine.

M.K. was 49 months when seen by the third author. At this time he had deficits in nonverbal communication, problems in peer relations, a lack of joint attention, poor conversational skills, stereotyped language and play, and preoccupations and routines. At this time he met the formal diagnostic criteria for Autistic Disorder. Regarding his language history, he showed no response to his name at 16 months. Single words emerged at 22 months, and two-word phrases at 27 months, although his language was echolalic. *What* and *Where* questions did not emerge until almost four years of age. Repetitive symbolic play emerged around 30 months. At 49 months, his Preschool Language Scale was in the average range, as was his Vineland Communication score. His Vineland Daily Living age equivalence was 3–0 and Socialization was 3–2. Spontaneous language was one-sided, tangential, scripted, and perseverative, and he still demonstrated some pronoun confusion. M.K. was diagnosed with Autistic Disorder on the basis of the following symptoms from the DSM-IV criteria: restricted nonverbal gestures, problems with eye gaze, difficulties developing peer relationships, lack of pointing to express interest, perseveration on topics of interest, stereotyped language, repetitive play, restricted and obsessive interests, and inflexible adherence to routines.

C.B. was 58 months when seen by the third author. At this time he had difficulties in nonverbal communication, peer relations, joint attention and emotional reciprocity. Conversational skills were delayed and he demonstrated stereotyped language and play, preoccupations and routines, and preoccupations with parts of objects. C.B. met the formal diagnostic criteria for Autistic Disorder at this time. Regarding his language history, he had a few single words by 18 months and some two-word phrases by 30 months. His speech at two years of age was echolalic and he did not ask questions until taught to do so at age three years. At 58 months C.B. still talked mainly on preferred topics and ignored conversational overtures by others. He had average performance on all formal language tests and Vineland Communication scores were within normal range, but Vineland Socialization and Daily Living scores were still delayed. C.B. was diagnosed with Autistic Disorder on the basis of the following symptoms from the DSM-IV criteria: restricted eye gaze, impaired nonverbal gestures, failure to develop peer relationships, lack of spontaneous pointing, impaired awareness of others, stereotyped language, lack of pretend play, adherence to routines, and preoccupation with parts of objects.

N.E. was 77 months when seen by the third author. He demonstrated continuing difficulties in peer relations, delayed language skills, impaired conversational abilities, and preoccupations and routines. N.E. met the formal diagnostic criteria for PDD-NOS at this time. Regarding his language history, N.E. said his first single words at 30 months and his first two-word phrases at 42 months. He often failed to respond to language and was echolalic at age three. An SLP evaluation at age six noted age appropriate language with maximal structure and support, but speech was often unintelligible and formulation was impaired. Testing by the third author at 77 months concurred and found language scores low average to average with particular problems with open-ended questions. Vineland Communication score age equivalence was 4–0, Daily Living was 2–9, and Socialization was 3–1. N.E. was diagnosed with PDD-NOS on the basis of the following symptoms from the DSM-IV criteria: inconsistent use of eye gaze, immature peer relationships, insensitive to others' distress, limited conversational skills, restricted patterns of interests, and inflexible adherence to routines.

L.X. was 56 months when seen by the third author. He continued to experience deficits in nonverbal communication, peer relations, and conversational skills. His language was delayed and stereotyped, and he had preoccupations and preoccupations with parts of objects. At this time he met the formal diagnostic criteria for Autistic Disorder. Regarding his language history, he had single words by 14 months, but his two-word phrases came in quite late. According to his mother, he was almost mute at age three. At age three-and-a-half, phrase speech showed notable echolalia. He followed one-step directions at age three, two-step directions at age four, used sentences and questions at age four, and his pretend play emerged at age four. At 56 months his Preschool Language Scores were above average on both the Receptive and Expressive subscales. His Vineland Communication score was within normal range, although he still did not volunteer information about experiences. Daily Living and Socializations subscales of the Vineland were both more than one year delayed. His language at this time showed a residual scripted quality and he still was unable to carry on a conversation. L.X. was diagnosed with Autistic Disorder on the basis of the following symptoms from the DSM-IV criteria: impaired eye gaze, lack of nonverbal gestures, difficulties in peer relationships, delayed language, limited conversational skills, stereotyped language, restricted pattern of interests, and preoccupation with parts of objects.

M.S. was 40 months when seen by the third author. At this time he had deficits in nonverbal communication, peer relations, joint attention, and emotional reciprocity. His language was delayed and conversational skills were a problem. He showed stereotyped language and play, as well as motor stereotypies and preoccupations with parts of objects. M.S. met the formal diagnostic criteria for Autistic

Disorder at this time. Regarding his language history, he did not babble as an infant and did not speak his first words until 33 months. He followed one-step directions at 36 months, two-step directions at 38 months, and pretend play emerged at 38 months. His emerging language was marked by flat prosody, echolalia, and scripts. He made rapid progress with ABA but at 40 months his language was still used only to protest and request and he showed no response to social language. At 40 months his vocabulary and grammar were within normal range on standardized tests although he still demonstrated difficulty with prepositions. Vineland Communication age-equivalence was 2–5, Daily Living 2–2, and Socialization 2–0. M.S. was diagnosed with Autistic Disorder on the basis of the following symptoms from the DSM-IV criteria: impaired eye-to-eye gaze, lack of nonverbal gestures, little interest in other children, lack of pointing to show interest, lack of emotional reciprocity, delayed language, impaired conversational skills, stereotyped language, restricted play, motor stereotypies, and preoccupation with parts of objects.

T.S. was 40 months when seen by the third author. He displayed deficits in nonverbal communication and peer relations, and had delayed language, preoccupations, routines, and preoccupations with parts of objects. At this time he met the formal diagnostic criteria for Autistic Disorder. Regarding his language history, he spoke his first single words at 17 months and first two-word phrases at 29 months. By thirty months response to sounds was inconsistent and deafness was questioned; humming and screaming were prevalent. He followed familiar one-step directions by two years, and two-step directions by three years. At 40 months he showed a thirteen-month receptive delay on the Preschool Language Scales and a fifteen-month expressive delay. Vineland Communication receptive age equivalence was 2–6, but his expressive age equivalence was 1–11. Socialization and Daily Living scores were in the low average range. T.C. was diagnosed with Autistic Disorder on the basis of the following symptoms from the DSM-IV criteria: delayed gestures, impaired eye gaze, difficulties in peer relationships, delayed language, delayed pretend play, stereotyped actions with objects, inflexible adherence to routines, and preoccupations with parts of objects.

T.M. was 79 months when seen by the third author. At this time he was still demonstrating deficits in nonverbal communication, peer relations, and conversational abilities. He had stereotyped play, preoccupations and routines, and motor stereotypies. T.M. met the formal diagnostic criteria for Autistic Disorder at this time. Regarding his language history, he spoke his first single words at 12 months and first sentences at two years. He did have some *wh*-questions at age three, but would still get them confused. Symbolic play and understanding two-step directions emerged at age

three. He had been diagnosed by a child neuropsychologist and a child psychiatrist as Asperger's due to poor pragmatic language and poor social competence. A recent SLP evaluation had shown average language except in the areas of listening, formulation, and pragmatics. Testing at 79 months showed all cognitive areas to be above average except processing speed, visuo-motor, and visuo-spatial skills. Vineland Communication score was average, Daily Living was low average, and Socialization showed a mild delay. T.M. was diagnosed with Autistic Disorder on the basis of the following symptoms from the DSM-IV criteria: limited nonverbal communication, lack of friends, lack of interest in peers, restricted conversational skills, restricted and stereotyped play, preoccupations and routines, inflexible adherence to routines, and stereotyped motor mannerisms.

U.N. was 71 months when seen by the third author. He was still experiencing mild delays in nonverbal communication and peer relations. He had delayed language and conversational skills, and demonstrated preoccupations and routines, and motor stereotypies. Formal diagnostic criteria for Autistic Disorder were not met because nonverbal communication and peer-relations deficits were regarded as being too mild; formal criteria for PDD-NOS were met. Regarding his language history, U.N. spoke his fist single words before 15 months but experienced a regression and age 15 months in social relatedness and a plateau in language development at 15–18 months, and showed no further gains in language development until he began treatment at age 27 months. At 71 months his Preschool Language Scale scores were within average range. Vineland Communication age-equivalence was 6–4, but Daily Living was 4–11 and Socialization was 4–2. U.N. was diagnosed with PDD-NOS on the basis of the following symptoms from the DSM-IV criteria: inconsistent eye contact, mild difficulties interacting with peers, restricted conversational skills, some stereotyped language, restricted patterns of interest, adherence to routines, and motor stereotypies.

T.B. was 56 months when seen by the third author. He demonstrated deficits in nonverbal communication, peer relations, and joint attention, and had delayed language and conversational skills. His language and play were stereotyped, and he experienced preoccupations with parts of objects. T.B. met formal diagnostic criteria for Autistic Disorder at this time. Regarding his language history, T.B. spoke his first single words at 28 months, his first phrases at 32 months, and his first sentences at 40 months. At age four-and-a-half, T.B. demonstrated prominent echolalia and pronoun reversal, and his expressive output was limited and scripted. At 56 months his language was still echolalic and scripted, with stilted prosody and pronoun reversal. His Preschool Language Scales Expressive age-equivalence was 3–0, and his receptive

age-equivalence was 3–6. The Vineland Communication age-equivalence was 3–5, Daily Living was 2–7, and Socialization was 2–9. T.B. was diagnosed with Autistic Disorder on the basis of the following symptoms from the DSM-IV criteria: impaired imitation, delayed nonverbal communication, lack of pointing to express interest, lack of social emotional reciprocity, lack of interest in other children, restricted play, difficulties with conversation, stereotyped language, and preoccupation with parts of objects.

Q.K. was 92 months when seen by the third author. He continued to experience problems in nonverbal communication, and had obsessive routines and motor stereotypies. He met the formal criteria for PDD-NOS. Regarding his language history, his early milestones were within normal range; however, his parents first became concerned at age three when he could not carry on a conversation, did not respond to questions, and was echolalic. Q.K. showed little or no joint attention or pretend play until he was taught at age five and his parents felt that he really did not understand language well until age six. Testing at 64 months shoed a VIQ of 55 and a PIQ of 99. Testing at 92 months showed listening comprehension low average and reading comprehension average. His Vineland Communication and Daily Living subscales were in the average range, with Socialization in the low average range. Q.K. was diagnosed with PDD-NOS on the basis of the following symptoms from the DSM-IV criteria: impaired eye gaze and nonverbal gestures, language delay, adherence to routine and motor stereotypies.

References

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

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a theory of mind? *Cognition, 21*, 37–46.

Bartolucci, G., Pierce, S. J., & Streiner, D. (1980). Cross-sectional studies of grammatical morphemes in autistic and mentally retarded children. *Journal of Autism and Developmental Disorders, 10*, 39–50.

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

Berman, R. A., Slobin, D. I., Stromquist, S., & Verhoeven, L. (1994). *Relating events in narrative: A crosslinguistic developmental study*. Mahwah, NJ: Lawrence Erlbaum.

Bibby, P., Eikeseth, S., Martin, N. T., Mudford, O., & Reeves, D. (2002). Progress and outcomes for children with autism receiving parent-managed intensive interventions. *Research in Developmental Disabilities, 23*, 81–104.

Cantwell, D., Baker, L., & Rutter, M. (1978). A comparative study of infantile autism and specific developmental receptive language disorder. IV. Analysis of syntax and language function. *Journal of Child Psychology and Psychiatry, 19*, 351–362.

Capps, L., Losh, M., & Thurber, C. (2000). ''The frog ate the bug and made his mouth sad'': Narrative competence in children with autism. *Journal of Abnormal Child Psychology, 28*, 193–204.

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

Chomsky, N. (1982). *Some concepts and consequences of the theory of government and binding*. Cambridge, MA: MIT Press.

deVilliers, J., & Roeper, T. (1995). Relative clauses are barriers to *wh*-movement for young children. *Journal of Child Language, 22*, 389–404.

Dunn, M., Gomes, H., & Sebastian, M. (1996). Prototypicality of responses in autistic, language-disordered, and normal children in a word fluency task. *Child Neuropsychology, 2*, 99–108.

Dunn, M., & Rapin, I. (1997). Communication in autistic children. In P. J. Accardo, B. K. Shapiro & A. J. Capute (Eds.), *Behavior belongs in the brain* (pp. 97–111). Baltimore, MD: York Press.

Fay, W. H., & Schuler, A. L. (1980). *Emerging language in autistic children*. Baltimore, MD: University Park Press.

Fein, D. (2001). The primacy of social and language deficits in autism. *Japanese Journal of Special Education, 38*, 1–16.

Fein, D., Dunn, M., Allen, D. A., Aram, D. M., Hall, N., Morris, R., & Wilson, B. C. (1996). Language and neuropsychological findings. In I. Rapin (Ed.), *Preschool children with inadequate communication* (pp. 123–154). London, UK: MacKeith Press.

Fenske, E. C., Zalenski, S., Krantz, P. J., & McClannahan, L. E. (1985). Age at intervention and treatment outcome for autistic children in a comprehensive intervention program. *Analysis and Intervention in Developmental Disabilities, 5*, 49–58.

Fowler, A. E. (1990). Language abilities in children with Down syndrome: Evidence for a specific syntactic delay. In D. Cicchetti & M. Beeghly (Eds.), *Down syndrome: The developmental perspective* (pp. 302–328). New York, NY: Cambridge University Press.

Fowler, A., Gelman, R., & Gleitman, L. R. (1994). The course of language learning in children with Down syndrome: Longitudinal and language level comparisons with young normally developing children. In H. Tager-Flusberg (Ed.), *Constraints on language acquisition: Studies of atypical children* (pp. 91–140). Hillsdale, NJ: Lawrence Erlbaum Associates.

Frith, U., & Happe, F. (1998). Why specific developmental disorders are not specific: On-line and developmental effects in autism and dyslexia. *Developmental Science, 1*, 267–272.

Gabriels, R. L., Hill, D. E., Pierce, R. A., Rogers, S. J., & Wehner, B. (2001). Predictors of treatment outcome in young children with autism. *Autism, 5*, 407–429.

Gardner, M. F. (1990). *Expressive one-word picture vocabulary test-revised*. Novato, CA: Academic Therapy.

Geller, E. (1998). An investigation of communication breakdowns and repairs in verbal autistic children. *The British Journal of Developmental Disabilities, 44*, 71–85.

Gelman, S. A., & Markman, E. M. (1986). Categories and induction in young children. *Cognition, 23*, 183–209.

Gelman, S. A., & Medin, D. L. (1993). What's so essential about essentialism? A different perspective on the interaction of perception, language, and conceptual knowledge. *Cognitive Development, 8*, 157–167.

Goldstein, H. (2002). Communication intervention for children with autism: A review of treatment efficiency. *Journal of Autism and Developmental Disorders, 32*, 373–396.

Happe, F. G. E. (1995). The role of age and verbal ability in the Theory of Mind task performance of subjects with autism. *Child Development, 66*, 843–855.

Harris, S. L., & Handleman, J. S. (2000). Age and IQ at intake as predictors of placement for young children with autism: A four-to six-year follow-up. *Journal of Autism and Developmental Disorders, 30*, 137–142.

Hauck, M., Fein, D., Maltby, N., Waterhouse, L., & Feinstein, C. (1998). Memory for faces in children with autism. *Child Neuropsychology, 4*, 187–198.

Hermelin, B., & O'Connor, N. (1970). *Psychological experiments with autistic children*. New York, NY: Pergamon Press.

Jackendoff, R. (1990). *Semantic structures*. Cambridge, MA: MIT Press.

Jocelyn, L. J., Casiro, O. G., Beattie, D., Bow, J., & Kneisz, J. (1998). Treatment of children with autism: A randomized control trial to evaluate a caregiver-based intervention program in community day-care centers. *Journal of Developmental and Behavioral Pediatrics*, 19, 326–334.

Johnson, S. C., & Carey, S. (1998). Knowledge enrichment and conceptual change in folkbiology: Evidence from Williams syndrome. *Cognitive Psychology, 37*, 156–200.

Kazak, S., Collis, G. M., & Lewis, V. (1997). Can young people with autism refer to knowledge states? Evidence from their understanding of ''know'' and ''guess''. *Journal of Child Psychology and Psychiatry, 38*, 1001–1009.

Kim, J., Marcus, G., Pinker, S., Hollander, M., & Coppola, M. (1994). Sensitivity of children's inflections to grammatical structure. *Journal of Child Language, 21*, 173–210.

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.

Levin, B. (1993) *English verb classes and alternations: A preliminary investigation*. Chicago: The University of Chicago Press.

Lord, C. (1996). Language in high-functioning adolescents with autism: Questions about deviance and delay. In D. Cicchetti & S. L. Toth (Eds.), *Adolescence: Opportunities and challenges. Rochester symposium on developmental psychopathology. Vol. 7* (pp. 149–165). Rochester, NY: University of Rochester Press.

Lord, C., & Paul, R. (1997). Language and communication in autism. In D. J. Cohen & F. R. Volkmar (Eds.), *Handbook of Autism and Pervasive Developmental Disorders*. New York, NY: Wiley Press.

Losh, M., & Capps, L. (2003). Narrative ability in high-functioning children with autism or Asperger's syndrome. *Journal of Autism and Developmental Disorders, 33*, 239–251.

Lovaas, I. O. (1987). Behavioral treatment and normal educational and intellectual functioning in young autistic children. *Journal of Consulting and Clinical Psychology, 55*, 3–9.

Lovaas, I. O., & Buch, G. (1997). Intensive behavioral treatment for young children with autism. In N. N. Singh (Ed.), *Prevention and treatment of severe behavior problems: Models and methods in developmental disabilities* (pp. 61–86). Pacific Grove, CA: Brooks/Cook Publishing.

Loveland, K. A., McEvoy, R. E., Tunali, B., & Kelley, M. L. (1990). Narrative story telling in autism and Down syndrome. *British Journal of Developmental Psychology, 8*, 9–23.

Mayer, M. (1969). *Frog, where are you?* New York, NY: Dial Press.

Mayeux, L. (2000). Linguistic, social, and cognitive factors in mental state verb and theory of mind development. Unpublished masters thesis, University of Connecticut, Storrs.

McEachin, J. J., Smith, T., & Lovaas, I. O. (1993). Long-term outcome for children with autism who received early intensive behavioral treatment. *American Journal on Mental Retardation, 97*, 359–372.

Menyuk, P., & Quill, K. (1985). Semantic problems in autistic children. In E. Schopler & G. B. Mesibov (Eds.), *Communication problems in autism* (pp. 59–68). New York, NY: Plenum Press.

Moore, C., Bryant, D., & Furrow, D. (1989). Mental terms and the development of certainty. *hild Development, 60*, 167–171.

Naigles, L. G., Fowler, A., & Helm, A. (1992). Developmental shifts in the construction of verb eanings. *Cognitive Development, 7*, 403–427.

Naigles, L., Fowler, A., & Helm, A. (1995) Syntactic bootstrapping from start to finish, with special reference to Down syndrome. In M. Tomasello & W. Merriman (Eds.), *Beyond names for things: Young children's acquisition of verbs* (pp. 299–330). Hillsdale, NJ: Lawrence Erlbaum Associates.

Naigles, L. G., Gleitman, H., & Gleitman, L. R. (1993). Children acquire word meaning from yntactic evidence. In E. Dromi (Ed.), *Language and cognition: A developmental perspective* (pp. 104–140). Norwood, NJ: Ablex Publishing.

Naigles, L. R., & Lehrer, N. (2002). Language-general and language-specific influences on children's acquisition of argument structure: A comparison of French and English. *Journal of Child Language, 29*, 545–556.

Perner, J., Leekham, S., & Wimmer, H. (1987). Three-year-olds difficulties with false belief. *British Journal of Developmental Psychology, 5*, 125–137.

Prizant, B. M. (1996). Brief report: Communication, language, social, and emotional development. *Journal of Autism and Developmental Disorders, 26*, 173–179.

Scarborough, H. S., Rescorla, L., Tager-Flusberg, H., Fowler, A. E., & Sudhalter, V. (1991). The relation of utterance length to grammatical complexity in normal and language-delayed groups. *Psycholinguistics, 12*, 23–45.

Shriberg, L. D., Paul, R., McSweeney, J. L., Klin, A., & Cohen, D. J. (2001). Speech and prosody characteristics of adolescents and adults with high-functioning autism and Asperger syndrome. *Journal of Speech, Language, and Hearing Research, 44*, 1097–1115.

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

Perner, J., Leekham, S., & Wimmer, H. (1987). Three-year-olds difficulties with false belief. *British Journal of Developmental Psychology, 5*, 125–137.

Szatmari, P., Bryson, S. E., Streiner, D. L., Wilson, F., Archer, L., & Ryerse, C. (2000). Two-year outcome of preschool children with autism or Asperger's syndrome. *American Journal of Psychiatry, 157*, 1980–1987.

Tager-Flusberg, H. (1985). Basic level and superordinate level categorization by autistic, mentally retarded, and normal children. *Journal of Experimental Child Psychology, 40*, 450–469.

Tager-Flusberg, H. (1988). On the nature of a language acquisition disorder: The example of autism. In F. S. Kessel (Ed.), *The development of language and language researchers: Essays in honor of Roger Brown* (pp. 249–267). Hillsdale, NJ: Lawrence Erlbaum Associates.

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.

Tager-Flusberg, H. (1996). Brief report: Current theory and research on language and communication in autism. *Journal of Autism and Developmental Disorders, 26*, 169–172.

Tager-Flusberg, H. (1999). A psychological approach to understanding the social and language impairments in autism. *International Review of Psychiatry, 11*, 325–334.

Tager-Flusberg, H. (2001). Understanding the language and communicative impairments in autism. *International Review of Research in Mental Retardation, 23*, 185–205.

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., & Sullivan, K. (1995). Attributing mental states to story characters: A comparison of narratives produced by autistic and mentally retarded individuals. *Applied Psycholinguistics, 16*, 241–256.

Thorndike, R. L., Hagel, E. P., & Satler, J. M. (1986). *Stanford-Binet intelligence scale. Fourth edition*. Chicago, IL: Riverside Publishing Co.

Waterhouse, L., & Fein, D. (1982). Language skills in developmentally disabled children. *Brain and Language, 15*, 307–333.

Wetherby, A. M., Prizant, B. M., & Schuler, A. L. (2000). Understanding the nature of communication and language impairments. In A. M. Wetherby & B. M. Prizant (Eds.), *Autism spectrum disorders: A transactional developmental perspective*. Baltimore, MD: Brookes Publishing.

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.

Wolk, L., & Giesen, J. (2000). A phonological investigation of four siblings with childhood autism. *Journal of Communication Disorders, 33*, 371–389.

Ziatas, K., Durkin, K., & Pratt, C. (1998). Belief term development in children with autism, Asperger syndrome, specific language impairment, and normal development: Links to theory of mind development. *Journal of Child Psychology and Psychiatry, 39*, 755–761.