Published in final edited form as:
J Autism Dev Disord. 2014 December ; 44(12): 3230–3244. doi:10.1007/s10803-014-2180-0.

Observation of Spontaneous Expressive Language (OSEL): A New Measure for Spontaneous and Expressive Language of Children with Autism Spectrum Disorders and Other Communication Disorders

So Hyun Kim, Ph.D.[1], **Dörte Junker, M.A.-CCC/SLP**[2], and **Catherine Lord, Ph.D.**[3]

[1]Yale Child Study Center, Yale University School of Medicine, 40 Temple St. Suite 7-D, New Haven, CT 06510, USA

[2]Milan Area Schools, 100 Big Red Drive, Milan, MI 48160, USA

[3]Center for Autism and the Developing Brain, Weill Cornell Medical College, 21 Bloomingdale Road, White Plains, NY 10605, USA

Abstract

A new language measure, the Observation of Spontaneous Expressive Language (OSEL), is intended to document spontaneous use of syntax, pragmatics, and semantics in 2-12-year-old children with ASD and other communication disorders with expressive language levels comparable to typical 2-5 year olds. Because the purpose of the OSEL is to provide developmental norms for use of language, the first step involves assessment of the scale's feasibility, validity, and reliability using a sample of 180 2-5 year-old typically developing children. Pilot data from the OSEL shows strong internal consistency, high reliabilities and validity. Once replicated with a large population-based sample and in special populations, the scale should be helpful in designing appropriate interventions for children with ASD and other communication disorders.

Keywords

Autism Spectrum Disorders; Language; Assessment

Communication impairments have been recognized as critical aspects of Autism Spectrum Disorder (ASD) from the first observations of children with the disorder (Kanner, 1943). Although the proportion of children with ASD who speak fluently has increased beyond earlier estimates of 50 percent of children as nonverbal (Lord, Risi, & Pickles, 2004), delayed language acquisition remains a hallmark of the disorder in early and middle childhood (Howlin, 2003) and language difficulties remain as key symptoms of ASD. These language deficits include pragmatic features such as the difficulty initiating and maintaining meaningful conversation (e.g. not responding to others' leads or questions; APA, 2013) and use of stereotyped speech or delayed echolalia (e.g. repeating lines from a Disney movie), as

Correspondence concerning this article should be addressed to So Hyun Kim, Yale Child Study Center, Yale University School of Medicine, 40 Temple St. Suite 7-D, New Haven, CT 06510, USA, so-hyun.kim@yale.edu; 203-764-5937; Fax: 203-764-4373. .

well as differences in use of even basic aspects of syntax such as difficulties in the use of grammatical morphemes (Bartolucci, Pierce, & Streiner, 1980; Eigsti, Bennetto, & Dadlani, 2007; Howlin, 1984; Roberts, Rice & Tager-Flusberg, 2004).

A valid assessment of communicative functioning in children with ASD beyond knowledge of vocabulary and syntax, in particular spoken language in more natural context, has significant implications for interventions. The emergence of spoken language in children with ASD is one of the most important variables predicting better outcomes in later childhood and adulthood (Gillberg & Steffenburg, 1987; Howlin et al., 2004; Venter, Lord, & Schopler, 1992). Improvement in communication is also one of the main goals in early treatments of ASD (Kasari, 2010; Smith, Groen, & Wynn, 2000). However, despite numerous treatment studies that have focused on language acquisition, lack of uniform measurement approaches to assessing language skills, especially use of language in less structured environments, has been problematic for the comparison of treatment outcomes across different intervention research (Tager-Flusberg et al., 2009).

Thus, a number of researchers have emphasized that assessments of spontaneous language are crucial in evaluating language deficits in children with specific language impairment as a complement to more structured measures (American Speech-Language-Hearing Association, 2008; Goffman & Leonard, 2000; Rescorla, Roberts, & Dahlsgaard, 1997). For instance, young children with ASD have been found to score higher on standardized language tests of vocabulary and verbal knowledge based on expressive language than they do on parent report focused on their ability to use language on a daily basis in natural settings (Luyster, Kadlec, Carter, & Tager-Flusberg, 2008). Though almost all models of language disorders include some attention to how children use language spontaneously in social contexts, in most structured language tests, spontaneous language (defined as language not directly elicited as part of the test) is not taken into account. For instance, existing measures elicit knowledge-based concepts by asking a child to label pictures or fill in blanks to assess syntactic or semantic skills (e.g., asking a child to look at a picture of boy running and answer, “Look at this boy; what is he doing?” to elicit use of a progressive verb; pointing to a picture of pigs and asking, “What are these?” to elicit use of a plural noun; as in the Preschool Language Scales [PLS, Zimmerman, Violette, Steiner, & Pond, 2011]). However, skills elicited in a highly structured manner may appear intact in children with ASD, even when they are significantly limited in their use of the same skills in more natural conversation. Several omnibus language tests such as the Reynell Developmental Language Scales (Reynell & Gruber, 1990) do require clinicians to obtain brief spontaneous language samples, but they typically occur in very constrained contexts (i.e. while looking at a picture or during a very brief activity with dolls). On the other hand, transcriptions or formal linguistic analyses (e.g. Child Language Data Exchange System [CHILDES], Systematic Analysis of Language Transcripts [SALT], Index of Productive Syntax [IPSyn; Scarborough, 1990]) can be used, but they are time consuming and require technical skills to score. Recognizing these needs, a new language assessment tool, the Observation of Spontaneous Expressive Language (OSEL), has been developed to measure children’s spontaneous use of language in standardized, but natural contexts.

Previous studies have also shown that, for many children, difficulties in social use of language far exceed what we might expect given their delays in other areas (Bishop 2002; Condouris, Meyer, & Tager-Flusberg, 2003; Tomblin et al., 2004). Children with ASD also demonstrate deficits in producing narrative discourse (e.g., causal statement; Tager-Flusberg, 1995; Losh & Capps, 2003). Loveland and colleagues (1990) found that children with ASD are also more likely to exhibit pragmatic violations including bizarre or inappropriate utterances while retelling a story compared to children with Down syndrome matched on chronological and verbal mental age.

Moreover, though some children with ASD score in the average or even higher ranges in standardized verbal IQ testing, which includes assessment of language skills (Kjelgaard & Tager-Flusberg, 2001; Tager-Flusberg, 2003), they still show significant abnormalities including odd use of language (e.g., stereotyped phrases, verbal rituals, pedantic speech) and limited reciprocity (e.g., not responding to others' conversational bids; limited ability ask questions about others' experiences; Ghazuddin, Tsai, & Ghazuddin, 1992; Gotham, Risi, Pickles, & Lord, 2006; Mayes, Volkmar, Hooks & Cicchetti, 1993; Paul, Orlovski, Marchinko, & Volkmar, 2009). Although some of traditional standardized measures attempt to assess pragmatic aspects of language in young children, these skills are typically elicited in limited contexts (i.e. asking the child to judge pragmatic rules of language or to use language appropriate in specific environmental situations depicted by pictures; such as in the Comprehensive Assessment of Spoken Language [CASL]; Carrow-Woolfolk, 2008).

Another approach is for skills to be assessed through checklists completed by a familiar person (e.g., caregiver, teacher, or clinician) as in the Children's Communication Checklist-2 [CCC-2]; Boshop, 2002, the Clinical Evaluation of Language Fundamentals-Preschool [CELFP]; Semel, Wiig, & Secord, 2004, and the Diagnostic Evaluation of Language Variation [DELV]; Seymour, Roeper, & de Villers, 2005). The Pragmatic Rating Scale-School Age (Landa et al., 1992) has also been used to document pragmatic skills in children with ASD ages 4 years through adulthood in conversational samples obtained in relatively non-standardized manners. Some of these measures assess subtle qualitative impairments in the use of language such as restricted and stereotyped conversational styles, abnormal prosody, or lack of coherent narrative, but others do not. To address these needs, the OSEL deliberately targets children's social use of language in standardized, but natural series of play contexts similar to the tasks used in the Autism Diagnostic Observation Schedule (ADOS; Lord et al., 2012) including narratives, the use of grammatical morphemes, semantic constructs, and various pragmatic skills.

Description of Instrument

The OSEL is a 30-45 minute observational assessment which focuses on children's spontaneous expressive language use in standardized, but natural contexts. The OSEL is intended to be used for children with ASD and other communication disorders from 2 years up to 12 years of age (depending on the interests of the child) whose language levels are equivalent to typically developing children from 2 to 5 years of age. The target population ranges from children who are beginning to use syntax (e.g., emerging 2- to 3-word phrases) to children who use complex sentences including 2 to 3 clauses. The OSEL is designed to be

used by a speech language pathologist, clinician, or researcher familiar with basic aspects of language structure and experienced in working with children with the disorders under consideration.

Using the general format of a range of unstructured to structured tasks similar to strategies in the ADOS, the examiner administering the OSEL presents highly motivating materials and activities organized in eight tasks that were specifically developed to elicit spontaneous use of expressive language in natural contexts (see Table 1). The first activity, *Play Figure Assembly*, gives the child the opportunity to interact within a relatively easy and usually familiar play context, namely to construct different play figures through adding body parts and accessories to plastic figures, in the case of bowling pins ("The Bowler Family"). The moderately low social and cognitive demand of the task provides a gentle warm-up activity for the child and creates opportunities for him/her to use different nouns, adjectives, and verbs as well as to answer and ask questions. The second activity, *Clarifications*, is incorporated anytime during the assessment and is focused on eliciting requests for clarifications from a child on the examiner's comments and questions. In order to do so, the examiner makes two comments and asks two questions to the child throughout different tasks. These comments and questions contain pseudo-words that most children are not familiar with so that the child can have opportunities to request for clarifications. This task can begin to be incorporated into the latter part of the Bowler Family, once the child is settled and resumed, as it fits, into other activities throughout the assessment. In the third activity, *Telling a Picture Story*, a child is presented with four pictures that depict a story or plot and then asked to describe the story to the examiner, thereby offering the opportunity for the child to demonstrate narrative skills. The fourth task, *Conversation*, may occur throughout the assessment as opportunities present themselves. This task requires the examiner to create contexts in which the child initiates conversations about points of interest or experiences and responds to conversational leads provided by the examiner. The fifth task, *Camping Trip/Picnic*, provides an opportunity to observe a child's spontaneous language production within a loosely structured, motivating pretend-play activity (e.g., getting ready for a camping trip or picnic, building a house/tent/castle, cooking and eating dinner). It also involves some exploratory/sensory activities (e.g., exploring with a flashlight, fishing on a pretend river made out of bubble wrap) in which a child has the opportunity to lead the interaction. Again, consistent with the other tasks, various morphological structures (e.g., verb phrases, verb tenses) as well as sentence structures (e.g., coordination, subordination), and pragmatic skills (e.g., asking for clarifications, making comments to express interests) can be observed during this task. The next task, the "*Where is it?" Game*, is intended to elicit the use of spatial prepositions. This task gives the child opportunities to verbally communicate locations of objects by using different prepositions. The seventh task, *Retell a Story: Where Are My French Fries?*, provides an opportunity for a child to re-tell a simple story that incorporates theory of mind. This gives the examiner a chance to observe the child's semantic and narrative skills (e.g., synthesizing information, understanding cause and effect relationships). In the last task, *Picture Description*, a child is shown a picture illustrating an open-ended story with some surprising elements (e.g., paint is about to spill on someone). The vignette provides opportunities for the child to describe objects, people, and events at multiple levels. The child is expected to use different word

classes (e.g., nouns, verbs, adjectives), and verb tenses (e.g., future, past, present) and forms (e.g., modal auxiliary) while describing the vignette.

The semi-structured, play-based environment of the OSEL was deliberately designed to evoke spontaneous use of a variety of syntax, pragmatics, and semantic structures based on children's spontaneous expressive language. During the administration of these different activities, the examiner codes an extensive list of syntactic skills and several pragmatic and semantic skills using a real time coding system. Additional pragmatic and semantic skills are coded after the administration using a Pragmatic Semantic Profile (PSP). All of these codes are transferred to a Summary Coding table after the administration.

Examples of items coded in the syntax coding sheet include the use of different verb forms including regular (e.g., He *climbed* up the tree.) and irregular past tenses (e.g., "I *caught* the fish!"), regular (e.g., I want *arms*.) and irregular plurals (e.g., "There are *geese* in the river."), and adjectives (e.g., "S'mores are my *favorite* snack."). Other more advanced syntactic skills are also coded in the OSEL (e.g., infinitive phrases, gerunds, negations, modal auxiliary verbs). In addition, throughout the OSEL administration, children are provided with opportunities to show various pragmatic skills including asking questions, offering information about their experiences, commenting on the materials and the examiner's actions, and clarifying what the examiner says. Other examples of pragmatic and semantic skills coded in the OSEL include reporting main ideas, synthesizing cause-and-effect information, and maintaining back-and-forth conversations. Besides these newly created items, the OSEL includes some modified codes from the ADOS that assess pragmatic skills and unusual features of language such as immediate echolalia and stereotyped/idiosyncratic use of words or phrases. These items have been expanded and elaborated for more detailed and comprehensive descriptions of these skills.

In order to assess these different aspects of children's expressive language, the examiner structures the OSEL tasks by adjusting what she or he says and the ways that materials are presented to create a context in which the child can use specific language skills spontaneously rather than the examiner deliberately eliciting them. For example, during the *Camping Trip/Picnic* task, the examiner follows a predetermined hierarchy of prompts to see if the child will request objects or actions. The hierarchy begins with the examiner waiting to see if the child initiates interaction when he wants the examiner to give him a toy fishing pole. If the child does not spontaneously request materials or activities, the examiner then looks deliberately at the child to see if he will say anything to request. Finally, if the child does not initiate a request, the examiner asks, "What would you like?" Thus, the OSEL coding reflects both how the child responded to the "press" for social behavior by using his/her pragmatic and semantic skills as well as how much the examiner had to structure the situation to elicit these responses.

This approach of using a predetermined hierarchy of prompts and structures in the administration of the OSEL is similar to that used in the ADOS. However, whereas the ADOS provides opportunities to elicit behaviors associated with core autism symptoms, the OSEL is designed to create opportunities to observe different aspects of expressive language skills which may or may not be associated with symptoms of ASD. As a result, the use of

the OSEL is not limited to children suspected of having ASD but also is intended for children who have specific language impairments (SLI), intellectual disabilities, or other developmental disabilities, and those who are suspected of having a language delay or disorder (beyond a phonological or articulation disorder) and are using language at less than a typically developing 5 year-old level and a chronological age from 2 to 12 years. Initial pilot testing of the OSEL took place informally with about 50 children with autism over several years as tasks and codes were developed and modified; however, the focus of this paper is on the performance of a systematically collected sample of typical children using the final version.

Aims

An ultimate goal of the OSEL is to obtain population norms from North America of standardized scores and age equivalents for each of the different areas (e.g., use of syntax, pragmatics, and semantics). However, before this standardization effort can begin, a proof-of-concept study that tests the extent to which the OSEL provides developmentally meaningful information about typical children is needed, as well as preliminary information about its feasibility and reliability with young children. This information is critical for researchers interested in using the instrument during the time that the full process of norm development is occurring. Therefore, the purpose of the present paper is to present data about the psychometric properties of the OSEL for an initial sample of 180 typically developing children from 2 to 5 years of age. The resulting data are not population norms, but do describe the performance of a preliminary comparison group to which we can compare the language skills of children with ASD or with other communication disorders whose language levels are comparable to that of typical 2 to 5 year olds. Based on the data for the typically developing children, preliminary age equivalent scores for syntax, pragmatics, and semantics were created to determine the usefulness of language profiles generated by the instrument. Eventually, data from a systematically determined sample of children as well as data specifically about different groups of children with ASD and/or language disorders will be collected, age equivalent scores will be revised, and standard scores will be created based on the larger population-based sample.

Methods

Participants

A sample was collected consisting of 180 typically developing children (95 males) between the ages of 2 to 5 years with a mean age of 41 months (*SD*=10.9) at the University of Michigan Autism and Communication Disorders Center (UMACC). A small subset of the sample (n=10) was reevaluated within 1-3 months to assess test-retest reliability. Another subset of the sample (n=10) was evaluated by the examiner and an observer during the live-administration to assess inter-rater reliability. All caregivers signed an Institutional Review Board approved informed consent to participate in research before participation. Children were divided into 7 groups by age (Table 2). The age range in each group was determined by the amount of age-related variance in the different factors (see below for a description of these). Participant characteristics including age, IQ scores, and scores on other language instruments are described in Table 2. Out of 180 children, 83% of children were Caucasian,

9% Biracial, 5% Asian, 2% African American, and 2% other races or unknown. All were native English speakers. 54% of mothers had degrees at a graduate or professional level, 28% from a four-year college, 17% at an associate or vocational level, and the rest were unknown.

Measures and Procedures

Both language and cognitive assessments were completed by a clinical psychologist, graduate student, or a trainee on the same day or within a few days' time. All of the data were collected by doctoral level students in clinical psychology and UMACC staff members who had achieved research reliability on the OSEL under the supervision of a licensed clinical psychologist (i.e., each examiner had to establish 80% agreement with a reliable coder on all items included in computing syntax and pragmatic-semantic totals for three consecutive administrations). For the 10 children for whom inter-rater reliabilities were assessed, in addition to the main examiner, another observer was present during the live-administration to code the child's language. For the cognitive testing, the Abbreviated Battery from the Stanford Binet Intelligence Scale (SBIS; Thorndike, Hagen & Sattler, 1986) was used to estimate verbal and nonverbal intellectual levels. To assess children's language level, all children were administered the Preschool Language Scale, Fourth Edition (PLS-4; Zimmerman, Steiner, & Pond, 2003) and/or the Comprehensive Assessment of Spoken Language (CASL; Carrow-Wodlfolk, 1999) depending on age and language level. Each caregiver completed the Vineland Adaptive Behavior Scales (VABS-II; Sparrow, Balla, & Cicchetti, 1984).

Coding on the OSEL

Grammatical uses for each syntax item were coded for different "types" of uses by excluding the grammatical uses of the same word or utterance. For example, the use of a question, "what's this," was coded only once for a WH- question even if the child used the same question again. Grammatical uses for 12 items were coded from 0 to 4 (e.g., copula verbs, modal auxiliaries); grammatical uses for two items were counted up to 3 because these particular items were used less frequently (e.g., progressive verbs, negations). Grammatical uses of 10 other items were counted up to ceilings of 6 to 18 because these particular items were more frequently used (e.g., regular past verbs, regular plurals) or consisted of sums of other items (e.g., sum of different types of subject pronouns). These 10 items were recoded into a 0- to 4-point scale with higher scores indicating more grammatical uses, no grammatical use coded as a score of 0, and remaining counts recoded to a score of 1, 2, 3, and 4 based on the original score distributions so that all items consistently fell on at 3-or 4-point scale.

The scores for the pragmatic and sematic profile (PSP) items ranged from 0 to 3 with higher scores indicating more impairment following the ADOS model. For example, for the item, *Verbal Requests to Get Needs Met*, a score of 0 meant that the child frequently used language to verbally request to get needs met. A score of 1 meant the child used language to verbal request but exhibits some instances in which the skill would have been expected and was not used. A score of 2 meant the child *occasionally* used language to verbally request but *consistently* exhibited instances in which the skill would have been expected and was not

used. A score of 3 meant that the child rarely or never requested verbally. A score of 3 for each item remained to capture the absence of a specific skill although it was converted to 2 to restrict the effects of any single item with extreme scores.

Statistical Analyses

In order to create total scores for syntax and pragmatic-semantic skills, after score distributions were examined, factor analyses were performed to examine the factor structure of all syntax and PSP items. The distribution of each syntax item was examined separately for different age and gender groups. The usefulness of each syntax item was assessed based on distributions of means, medians, and standard deviations. Similarly, we examined means, medians, and standard deviations for item scores in the pragmatic semantic profile (PSP) separately by gender and age groups. As a next step, the factor structure of all syntax items used to generate the OSEL syntax totals was examined by performing exploratory factor analysis (EFA) using Mplus software (Muthen & Muthen, 1998). An EFA was also performed to examine the factor structure of all PSP items. For PSP only, for which more than one factor emerged from EFA, a confirmatory factor analysis (CFA) was then performed. Based on the factors emerging from these analyses, the PSP subdomain totals were calculated by adding the scores of items within each factor.

Reliability and validity of the OSEL were assessed based on the OSEL syntax and PSP totals. Intraclass correlations (ICC) were calculated for inter-rater and test-retest reliabilities using the OSEL syntax and PSP item and total scores obtained from a subset of samples (n=10). Weighted kappas (Cohen, 1960) were also calculated for inter-rater and test-retest reliabilities based on all items except those with very unbalanced ratings for which percent agreements were reported alternatively (e.g., for the item, *response to yes/no questions*, 9 out of 10 subjects were rated with a score of 3 resulting in a high agreement but very unbalanced rating, yielding a non-interpretable result; Cicchetti & Feinstein, 1990). Concurrent validity was assessed through calculating correlations between the OSEL syntax and PSP totals and scores from other measures such as the Expressive Language subdomain age equivalents in a measure of adaptive behavior (VABS) and domain standard scores from other language measures (PLS-4 Expressive Communication and Auditory Comprehension, CASL Syntax Construction and Pragmatic Judgment). The relations between the OSEL scores and chronological age and verbal and nonverbal IQ scores (SBIS) were also examined as indicators of convergent validity and specificity of the OSEL. In addition, internal consistency of all items was examined by calculating Cronbach's alpha for all participants for each age group separately for syntax and PSP items. Finally, the effects of gender, age, and verbal IQ on the OSEL syntax and PSP totals were examined using General Linear Models.

Although we are aware that there are limitations in the use of age equivalents, we felt that representing scores in this preliminary version as age equivalents best matched the developmental intentions of the scale (as opposed to a goal of classifying children by the number of deviations from the mean that their scores indicated). Age equivalents can also be particularly valuable to caregivers, educators, and therapists in interpreting language profiles, within the context of acknowledging their imprecision. Thus, preliminary age

equivalents were derived for the OSEL syntax and PSP totals separately by gender using a standard method (Ward, Stoker, & Murray-Ward, 1996). First, median scores from the syntax and PSP totals were computed separately for each age group and gender. These median scores were plotted across age groups. A smooth line was fitted to the plotted points. Age equivalents corresponding to particular syntax and PSP totals were then read from the smooth lines separately by gender based on the equations of the lines. Thus, the age equivalents represent ages corresponding to median syntax and PSP scores predicted by the fitted line. This is consistent with other commonly used language measures for which age equivalents were developed (e.g., CASL [Carrow-Wodlfolk, 1999]; PLS-4 [Zimmerman, Steiner, & Pond, 2003]). The fit of the line was high in ranging from a R^2 of 0.82 to 0.98 for syntax and PSP totals by gender, which also supports the validity of the measure as a developmental tool.

Results

Feasibility of the OSEL

All 4 examiners achieved reliability within a 2-week-training during which each examiner administered and coded 3-5 practice cases. Out of the 4 examiners, three did not have a previous background in speech and language evaluations although they had general backgrounds in psychology. The examiners were able to complete all components of the assessment within 30-45 minutes for all children who participated in the study. This included the “live” coding of the syntax items as well as several pragmatic-semantic skills. The examiners spent 10-15 additional minutes after the administration to complete the PSP and finalize the coding. Therefore, in general, the examiner spent approximately 1 hour administering and coding the OSEL.

Creating Syntax and PSP Totals and Domain Totals based on Examination of Score Distribution and Factor Analyses

A total of 24 out of 28 items on either 3-point or 4-point scale that showed gradual increases in means and medians from younger to older ages were included in a final set of items to create the “OSEL syntax totals.” The OSEL syntax totals were intended to reflect the developmental progression of syntactic skills across development. Thus, two items (declarative and imperative sentences) were excluded from the final set of analyses because they were equally prevalent in all age groups. Two other items added toward the end of data collection (WH- infinite phrases and gerunds) were also excluded from these analyses because there was not sufficient data to be analyzed. A total of 15 items in the PSP on the 3-point scale were combined to create the “OSEL PSP totals.” Three items added toward the end of data collection (level of support required for conversation, intonation/volume/rhythm/rate, intelligibility) are not included in OSEL PSP totals here because of limited data.

Results from the exploratory factor analysis (EFA) using the OSEL syntax items showed that a 1-factor solution fitted well (Table 3) with the goodness-of-fit rating of a Comparative Fit Index (CFI) of 0.99 (CFI between 0.9 and 1 indicating good fit; Skrondal & Rabe-Hesketh, 2004) and a Root Mean Square Error Approximation (RMSEA) of 0.057 (RMSEA of 0.08 or less is considered a satisfactory fit; Browne & Cudeck, 1993). Notably, the OSEL

syntax totals reached the ceiling for girls at around 48-53 months whereas for boys, the scores continued to increase until 5 years of age.

Results from the EFA using the OSEL PSP items showed that a 3-factor solution fitted well (Table 4). Items loaded onto three factors (See Table 4 for the item loadings). The first factor appeared to represent skills related to *Initiation of Reciprocal Communication*, which included items such as Verbal requites to get needs met, Asks for information about thoughts, feelings, or experiences, and Comments or offers information about thoughts, feelings, or experiences. The second factor included loadings from items such as Reporting main ideas, Reporting sequence of events/story, and Synthesizing cause-and-effect information, representing *Narrative Skills*. The third factor was consistent with *Unusual Features*, including Interrupts the examiner or dominates conversations, Stereotyped/ idiosyncratic use of words or phrases, and *Unspecific language and/or semantic errors*.

One of the items, *Stereotyped language*, was excluded from the EFA due to the large portion of children scoring 0s (more than 90% of children in this typical sample). The item was included for the CFA because it is anticipated that many more children with ASD will have scores other than 0 on this item. Another item, *Repairs and request for clarification* loaded higher on the third factor, *Unusual Features*, was left under the second factor, *Narrative Skills*, for theoretical reasons, and when the item was included in the second factor for the CFA, it showed a high loading with the second factor. The goodness-of-fit rating yielded a CFI of 0.995 and 0.978 and a RMSEA of 0.048 and 0.080 for EFA and CFA respectively. Based on the three factors emerging from the analyses, PSP subdomain totals were calculated by combining item scores under each domain. “PSP 3 domain totals” were also created by adding item scores under all three domains. The mean syntax and PSP totals by gender and age groups are presented in Table 5.

Reliabilities

The sample used for inter-rater reliabilities consisted of 10 children (6 males) whose mean age was 40.4 months (*SD*=12.4) with a mean syntax total of 59.6 (*SD*=18.8) and a mean PSP total of 11.8 (*SD*=3.8). Intraclass correlation (ICC) between raters was 0.96 for the syntax totals and 0.83 for the PSP totals (both $p<0.001$). Weighted kappas for inter-rater reliabilities for the syntax totals ranged from 0.54 to 1 (all $p<0.01$; $K_w=0.54$ for *Object Pronouns*; $K_w=1$ for *Future Tense Verbs*, *Nouns*, and *Verbs*; Cohen’s kappa between 0.21 to 0.40 is considered as “Fair,” 0.41 to 0.60 as “Moderate,” and 0.61 to 0.80 as “Substantial,” 0.81 to 1.00 as “Almost perfect;” Landis and Koch, 1977). Most weighted kappas for the PSP totals were significant ranging from 0.41 to 1 (all $p\leq0.05$; $K_w=0.41$ for *Repairs/Request clarification*; $K_w=1$ for Interrupts the examiner or dominates conversations, Stereotyped/ Idiosyncratic use of words or phrases, Unspecific language and/or semantic errors, and *Impolite or inappropriate language*) except for one item, Comments on characters’ emotional and/or mental states ($K_w=0.33$, $p=0.098$).

The sample used for test-retest reliabilities consisted of another set of 10 children (7 males) whose mean age was 45.6 months (*SD*=10.1) with a mean syntax total of 74.6 (*SD*=6.5) and a mean PSP total of 6.7 (*SD*=1.9). ICC for test-retest reliabilities was 0.95 for the syntax totals and 0.92 for the PSP totals (both $p<0.001$). Weighted kappas varied more for test-

retest than for inter-rater reliabilities. Most kappas were significant ranging from 0.35 to 1 for the majority of syntax items (all $p<0.05$; $K_w=0.35$ for *Progressive Tense Verbs*; $K_w=1$ for R*esponses To Yes/No Questions* and *Verbs*) with a few exceptions such as *Coordination, Infinitive Phrases*, and *Negation* ($K_w=0.30$, $K_w=0.25$, $K_w=0.19$, respectively; p=n.s.). Most weighted kappas for the PSP items were also significant ranging from 0.3 to 1 (all $p\leq0.05$; $K_w=0.3$ for Asks for information about thoughts, feelings, or experiences; $K_w=1$ for *Maintains a conversation*, *Stereotyped/Idiosyncratic use of words or phrases,* and *Unspecific language and/or semantic errors*) except for one item, Comments on characters' emotional and/or mental states ($K_w=0.14$, $p=0.245$). See Supplemental Table 1 for more details.

Concurrent and Convergent Validity

As expected, the Pearson r correlation between the OSEL syntax totals and chronological age was relatively high ($r=0.6$, $p<0.01$). The OSEL syntax totals were also correlated with the CASL Syntax Construction domain standard scores ($r=0.6$, $p<0.01$) and the CASL Pragmatic Judgment domain standard scores ($r=0.5$, $p<0.01$) using a subset of 112 children who were old enough to be given the CASL. Across all age groups, the correlation between the OSEL syntax totals with the PLS-4 Expressive Communication domain scores was 0.4 ($p<0.01$). The correlation between the OSEL syntax totals and the PLS-4 Auditory Comprehension domain scores was also 0.4 ($p<0.01$) for all participants. The correlation between the OSEL and the VABS Expressive Language subdomain was 0.7 ($p<0.01$). Correlations between the OSEL scores and estimated verbal and nonverbal IQ scores were r of 0.3 ($p<0.01$) for both verbal and nonverbal IQ scores.

The OSEL PSP 3 domain totals were also relatively highly correlated with age ($r=-0.6$, $p<0.01$). Across all age groups, the correlation between the OSEL PSP 3 domain totals (combined scores of items in all three domains; higher scores indicating absence/abnormality of skills specified) and the PLS Expressive Language was -0.4 ($p<0.01$). The correlation between the OSEL PSP 3 domain totals and the PLS Auditory Comprehension domain scores was -0.4 ($p<0.01$). The OSEL PSP 3 domain totals were also correlated with the CASL Syntax Construction and Pragmatic Judgment standard scores (both $r=-0.5$, $p<0.01$, n=112). The correlation with the VABS Expressive Language subdomain was -0.6 ($p<0.01$). Correlations between the OSEL PSP 3 domain totals with verbal and nonverbal IQ scores were both -0.3, ($p<0.01$).

All of these findings were as expected, confirming that the OSEL scores depicted a developmental progression associated with age and other measures of language skills in typically developing children, but with less than perfect correlations, suggesting that we were measuring concepts that were related but not identical to standard measures of language ability (PLS-4; CASL), verbal IQ (SBIS), and everyday communication (VABS).

Internal Consistency

For all syntax items, Cronbach's alpha ranged from 0.84 to 0.92 for different age groups with the Cronbach's alpha across all age groups at 0.938 ($p<0.001$). For all PSP items, Cronbach's alpha ranged from 0.68 to 0.8 with the Cronbach's alpha across all age groups at

0.8 for the PSP items ($p<0.001$). Lower alphas were expected for the PSP items since three different constructs were combined into one single score as indicated by the factor analyses.

Effects of Gender, Age, and Verbal IQ as Predictors of OSEL Syntax and PSP Totals

The General Linear Model showed that gender was a significant predictor of the OSEL syntax totals ($F=5.55$, $p<0.05$) and for the PSP *Initiation of Reciprocal Communication* domain totals ($F=5.14$, $p<0.05$) while controlling for age and verbal IQ. Age significantly predicted the syntax totals ($F=212.97$, $p<0.001$) and all PSP totals ($F=94.76$ for *Initiation of the Reciprocal Communication* domain totals, $F=135.35$ for the *Narrative Skills* domain totals, $F=28.79$ for the *Unusual Features* domain totals, $F=185.16$ for the PSP 3 domain totals, all $p<0.001$). Verbal IQ was a significant predictor of the syntax totals ($F=53.63$, $p<0.001$) and all PSP totals ($F=17.11$ for the *Initiation of Reciprocal Communication* domain totals, $F=19.91$ for the *Narrative Skills* domain totals, $F=19.5$ for the *Unusual Features* domain totals, $F=39.18$ for the PSP 3 domain totals, all *p's*<0.001).

Deriving Preliminary Age Equivalents for Syntax and Pragmatic Semantic Profile Totals

The fit for the smooth lines based on the median syntax totals across age groups was R^2 of 0.9 for males and 0.94 for females. The fit for the PSP totals ranged from R^2 of 0.85 to 0.95 for males and from 0.85 to 0.99 for females by different factors. Figure 1 shows an example of the smooth line fitted for the medians of syntax totals for males. Age equivalents calculated from the smooth lines for the OSEL syntax and PSP totals for males and females are presented in Figure 2 (Ward, Stoker, & Murray-Ward, 1996). Because the behaviors specified under the PSP items that loaded onto the *Unusual Features* factor (e.g., Stereotyped/idiosyncratic use of words or phrases, Immediate echolalia) were rare in the typically developing children included in the current sample, scores for these items were relatively low. The mean totals for this factor ranged from 0.1 to 2.3 with standard deviations ranging from 0.3 to 2.3 (See Table 5). Thus, age equivalents were not created for this factor due to the limited variability across age groups. However, item scores from this domain were included in the OSEL PSP 3 factor totals.

Discussion

The OSEL is a measure of children's spontaneous expressive language obtained in standardized, but natural contexts. Results indicate the sensitivity of the OSEL to developmental changes in use of syntax, semantics, and pragmatics in typical children across the age range of 2 to 5 years. We were also able to demonstrate the feasibility of using the OSEL tasks with children as young as 2 years and of using the codes to describe the language of children up to 4 and a half years for girls and 5 for boys, of the examiners doing live coding of syntax, and of pragmatic and semantic codes immediately following the session. Thus, in about an hour, the OSEL successfully captured different aspects of expressive language skills (i.e., syntax, pragmatics, and semantics) in the preliminary sample of typically developing 2-to-5-year-olds.

Results based on the typical sample also indicate excellent internal consistency for the OSEL syntax and pragmatic-semantic items as well as moderate to high inter-rater

reliabilities and fair to high test-retest reliabilities for most items. Concurrent validity was observed through moderate to strong associations between the OSEL syntax and PSP totals and other language measures (e.g. Expressive and Receptive domains from the PLS-4 and Pragmatic Judgment and Syntax Constructions subtests from the CASL, and VABS Expressive Language subdomain). Significant, moderate correlations between the OSEL and other instruments were expected because the OSEL measures language skills related but not identical to the linguistic knowledge and adaptive expressive language skills targeted in well-established tests such as the PLS, CASL, and VABS.

Results from the general linear regression analysis showed that gender made a significant independent contribution to the OSEL syntax and PSP totals. Consistent with past research suggesting that language acquisition is more rapid for females than for males during toddler and early preschool years (Bauer et al., 2002; Galsworthy, Dionne, Dale, & Plomin, 2000), females showed significantly more grammatical uses and advanced pragmatic and semantic skills than males across all ages on the OSEL, with the gap between males and females decreasing over time. As a result, age equivalents were created separately by gender.

The correlations between verbal IQ and the OSEL syntax and PSP totals were minimal remaining at *r* of 0.3. However, the general linear regression analysis also showed that verbal IQ scores made significant independent contributions to the OSEL scores. Although the OSEL emphasizes different aspects of language than verbal IQ, the OSEL scores are not expected to be completely independent of IQ given the role of language skills in the measurement of cognitive skills in young children. In fact, some of the items in the Abbreviated Battery from the Stanford Binet Intelligence Scale have been found to be highly associated with language skills (Thorndike, Hagen & Sattler, 1986).

One of the advantages of the OSEL is that it focuses on children's spontaneous use of expressive language in standardized, but natural contexts (e.g. while playing with a variety of toys, telling stories from a picture vignette, and interacting with an examiner during imaginative play). This is different from most standardized testing, which elicits responses that are knowledge-based or highly tied to concepts (e.g., This chain is long, this chain is…) rather than spontaneous expressive skills. By using various play-based tasks in the OSEL, researchers and clinicians can obtain meaningful profiles of spontaneous language skills that are more reflective of how children use their language skills in everyday activities (e.g., at home while interacting with parents and siblings, at school while interacting with teachers and peers) and compare them to language abilities measured in a more structured test. Both types of information are important and need to be used together to determine whether treatment is appropriate, if so, what should be done, and what the goals are.

Another advantage of using the OSEL is that researchers and clinicians can obtain quantified profiles of spontaneous expressive language for children with ASD and other communication disorders for syntactic, pragmatic, and semantic skills. The quantified language profiles obtained from the OSEL offer an important complement to the well-established standardized language tests that emphasize knowledge of language forms more than spontaneous use. Most of these skills are coded "live" using online coding systems, which is a unique feature of the OSEL. Video recording of the sessions can also be used for

future investigation and research. More importantly, the quantified and detailed information obtained in the distinctive areas of language can help clinicians develop individualized intervention programs and provide quantifiable justification for services for children with ASD and other communication disorders who need support in using language, even when their knowledge of linguistic concepts may be adequate. The OSEL can also provide a standardized method to measure changes in expressive language skills over the course of treatment.

Limitations and Strengths

The sample included in the present study has clear limitations. A large proportion (more than a half) of the children had parents with relatively high educational backgrounds (graduate or professional level education). The mean scores for the PLS for this typical sample were above average even though scores for the other instruments were not as high. The current data was not based on a population sample, and was limited to 180 children. The reliability analyses were based on a small number of cases; all results must be confirmed with a larger sample. Age equivalents derived from the present study are preliminary and will eventually need to be replicated with a larger, more representative sample. With a larger normative sample, distributions of item scores that were added toward the end of the data collection process (gerunds and conjunctions for syntax items; level of support required for conversation, intonation/volume/rhythm/rate, and intelligibility for PSP items) can be examined further to test the feasibility of including these items in the total scores and age equivalents.

Because the OSEL was validated with typically developing children, raw item scores under the PSP *Unusual Features* domain, originally created for clinical populations, were lower than the scores on the other two domains. Thus, the item score distributions under this domain (e.g., Stereotyped/idiosyncratic use of words or phrases, Immediate echolalia, Semantic errors) should be reexamined in clinical populations. Based on the earlier pilot work with children with ASD, it is expected that children with ASD will have significantly higher scores than children in the current study on these items. Systematic research is needed with clinical populations to identify the pattern of language impairments in this area.

This paper only describes data from typical children. It is not within the scope of a single paper to present data for children with ASD or other disorders. The tasks in the OSEL were designed initially for children with ASD, working generally from the strategies used in the ADOS; this paper shows that these strategies resulted in the ability to document developmental progressions in syntax, semantics, and pragmatics in typical children from 2 to 4-5 years of age.

Despite these limitations, the current study has a number of strengths. This is one of the first studies attempting to quantify spontaneous language skills that are easily missed by formal language tests. Although the sample is restricted to typically developing children from 2-5 years, results show strong feasibly of the OSEL for the use of the instrument to measure pragmatic and semantic skills that are rarely addressed by other instruments at this age range. Since the OSEL has been developed based on a range of developmental and language levels, the continuity in scoring from beginning syntax to more complex levels of discourse

as well as from young preschoolers to elementary school-aged children will allow comparisons of children across different chronological ages, language levels, and genders. Once these results are replicated with other population-based and clinical samples, the OSEL will be particularly useful for young preschoolers with ASD and other developmental disorders who have relatively mild language delays, as well as older, school-aged children with moderate communication impairments.

Conclusion

The OSEL is a measure of children's spontaneous use of language in standardized, but natural contexts. In a relatively brief time period (less than an hour), the OSEL provides quantified profiles of spontaneous expressive language use in typically developing children from 2 to 4 to 5 years of age using syntax and pragmatic-semantic totals and age equivalents. It is hoped that the OSEL can be used in combination with other language measures to evaluate strengths and weaknesses of expressive language skills in children with ASD and other communication disorders from 2 up to about 12 years (Tager-Flusberg et al., 2009). In the near future, using a sample of children with ASD and other developmental disorders (e.g., language delays, intellectual disabilities), the validity of the measure will be further evaluated by comparing the distributions of item scores across different diagnostic categories. Children with ASD and other communication disorders would be expected to show impairment in morphosyntactic skills when compared to typically developing children. It is also expected that children with ASD will show more difficulty in pragmatic and semantic skills compared to children with other communication disorders and/or typically developing children. In the end, the ultimate goal is to generate an individualized profile for children with ASD or other communication disorders. Based on a nationally representative standardization sample from an ongoing large-scale study, age equivalents will be replicated and standard scores for syntax, semantics, and pragmatics will be created.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was supported by grants from the National Institute on Deafness and Other Communication Disorders (1R21DC007921-01) and the Western Psychological Services. We gratefully acknowledge the children and their families who participated in the study, and the faculty and staff at the University of Michigan Autism and Communication Disorders Center (UMACC) and the Center for Autism and the Developing Brain (CADB) and our collaborators, particularly Kate Gasparrini, Kristin Houck, Alayna Schreier, Sheri Stegall, Jacquelyn Baker, Rebecca Jones, Daniel Cheong, Vanessa Hus, and Shanping Qiu who assisted in collecting, preparing, analyzing, and interpreting these data.

References

American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 5th ed. Author; Washington, DC: 2013.

American Speech-Language-Hearing Association. [Retrieved January 24, 2013] Roles and responsibilities of speech-language pathologists in early intervention: Guidelines. ASHA Practice Policy. 2008. from http://www.asha.org/policy/GL2008-00293.htm

Alarcon M, Canton RM, Liu J, Gilliam TC, Geschwind DH, Autism Genetic Research Exchange Consortium. Evidence for a language quantitative trait locus on chromosome 7q in multiplex autism families. American Journal of Human Genetics. 2002; 70:60–71. [PubMed: 11741194]

Barrett S, Beck JC, Bernier R, Bisson E, Braun TA, Casavant TL, Childress D, et al. An autosomal genomic screen for autism: Collaborative linkage study of autism. American Journal of Medical Genetics. 1999; 88:609–615. [PubMed: 10581478]

Bartolucci G, Pierce SJ, Streiner D. Cross-sectional studies of grammatical morphemes in autistic and mentally retarded children. Journal of Autism and Developmental Disorders. 1980; 10(1):39–50. [PubMed: 6927677]

Bauer DJ, Goldfield BA, Reznick JS. Alternative approaches to analyzing individual differences in the rate of early vocabulary development. Applied Psycholinguistics. 2002; 23:313–335.

Bishop DVM. Putting language genes in perspective. Trends in Genetics. 2002; 18(2):57–59. [PubMed: 11818129]

Bishop, DVM. The Children's Communication Checklist version 2 (CCC-2). Psychological Corporation; London: 2003.

Bishop DVM, Norbury CF. Exploring the borderlands of autistic disorder and specific language impairment: a study using standardized diagnostic instruments. Journal of Child Psychology and Psychiatry. 2002; 43(7):917–929. [PubMed: 12405479]

Bradford Y, Haines J, Hutcheson H, Gardiner M, Braun T, Sheffield V, et al. Incorporating language phenotypes strengthens evidence of linkage to autism. American Journal of Medical Genetics. 2001; 105(6):539–47. [PubMed: 11496372]

Brown, R. A first language: The early stages. Harvard U. Press; Oxford England: 1973.

Cicchetti D, Feinstein A. High agreement but low kappa: II. Resolving the paradoxes. Journal of Clinical Epidemiology. 1990; 43(6):551–558. [PubMed: 2189948]

Cicchetti D, Volkmar F, Klin A, Showalter D. Diagnosing autism using ICD-10 criteria: A comparison of neural networks and standard multivariate procedures. Child Neuropsychology. 1995; 1:26–37.

Cohen, Jacob. A coefficient of agreement for nominal scales. Educational and Psychological Measurement. 1960; 20(1):37–46.

Condouris K, Meyer E, Tager-Flusberg H. The relationship between standardized measures of language and measures of spontaneous speech in children with autism. American Journal of Speech-Language Pathology. 2003; 12:3–15. [PubMed: 12680809]

Eigsti I, Bennetto L, Dadlani MB. Beyond pragmatics: Morphsyntactic development in autism. Journal of Autism and Developmental Disorders. 2007; 37:1007–1023. [PubMed: 17089196]

Farrell AD, Mariotto MJ, Conger AJ, Curran JP, Wallander JL. Self-Ratings and judges' ratings of heterosexual social anxiety and skill: A generalizability study. Journal of Consulting and Clinical Psychology. 1979; 47:164–175. [PubMed: 429650]

Fenson L, Pethick SJ, Renda C, Cox JL, Dale PS, Reznick JS. Short-form versions of the MacArthur communicative development inventories. Applied Psycholinguistics. 2000; 21:95–115.

Galsworthy MJ, Dionne G, Dale PS, Plomin R. Sex differences in early verbal and non-verbal cognitive development. Developmental Science. 2000; 3:206–215.

Ghaziuddin M, Tsai L, Ghaziuddin N. Brief Report: A comparison of the diagnostic criteria for Asperger syndrome. Journal of Autism and Developmental Disorders. 1992; 22:643–649. [PubMed: 1483981]

Goffman L, Leonard J. Growth of language skills in preschool children with specific language impairment: Implications for assessment and intervention. American Journal of Speech-Language Pathology. 2000; 9:151–161.

Gotham K, Risi S, Pickles A, Lord C. The Autism Diagnostic Observation Schedule: revised algorithms for improved diagnostic validity. Journal of Autism and Developmental Disorders. 2007; 37(4):613–627. [PubMed: 17180459]

Gillberg C, Steffenburg S. Outcome and prognostic factors in infantile autism and similar conditions: A population-based study of 46 cases followed through puberty. Journal of Autism and Developmental Disorders. 1987; 17(2):273–287. [PubMed: 3610999]

Harris G, Chabris C, Clark J, Urban T, Aharon I, Steele S, et al. Brain activation during semantic processing in autism spectrum disorders via functional magnetic resonance imaging. Brain and Cognition. 2006; 61:54–68. [PubMed: 16473449]

Howlin P. The acquisition of grammatical morphemes in autistic children: A critique and replication of the findings of Bartolucci, Pierce, and Streiner, 1980. Journal of Autism and Developmental Disorders. 1984; 14(2):127–136. [PubMed: 6746504]

Howlin P. Outcome in high-functioning adults with autism with and without early language delays: implications for the differentiation between autism and Asperger syndrome. Journal of Autism and Development Disorders. 2003; 33:3–13.

Howlin P, Goode S, Hutton J, Rutter M. Adult outcome for children with autism. Journal of Child Psychology and Psychiatry. 2004; 45(2):212–229. [PubMed: 14982237]

Kasari C, Gulsrud A, Wong C, Kwon S, Locke J. Randomized controlled caregiver mediated joint engagement intervention for toddlers with autism. Journal of Autism and Developmental Disorders. 2010; 40:1045–56. [PubMed: 20145986]

Kanner L. Autistic disturbances of affective contact. Nervous Child. 1943; 2:217–250.

Kjelgaard M, Tager-Flusberg H. An investigation of language impairment in autism: Implications for genetic subgroups. Language and Cognitive Processes. 2001; 16:287–308. [PubMed: 16703115]

Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977; 33(1):159–174. [PubMed: 843571]

Landa R, Piven J, Wzorek M, Gayle J, Chase G, Folstein S. Social language use in parents of autistic individuals. Psychological Medicine. 1992; 22:245–254. [PubMed: 1574562]

Leonard, L. Children with Specific Language Impairment. The MIT Press; Cambridge, MA: 1998.

Lord, C.; Rutter, M.; DiLavore, PC.; Risi, S.; Gotham, K.; Bishop, S. The Autism Diagnostic Observation Schedule, Second Edition (ADOS-2) Manual (Part 1): Modules 1-4. Western Psychological Services; Torrance, CA: 2012.

Lord, C.; Risi, S.; Pickles, A. Trajectory of language development in autistic spectrum disorders. In: Warren, SF., editor. Developmental Language Disorders: From Phenotypes to Etiologies. Lawrence Erlbaum Associates Publishers; Mahwah, NJ US: 2004. p. 7-29.

Lord C, Schulman C, DiLavore P. Regression and word loss in autistic spectrum disorders. Journal of Child Psychology and Psychiatry. 2004; 45:936–955. [PubMed: 15225337]

Losh M, Capps L. Narrative ability in high-functioning children with autism or Asperger's syndrome. Journal of Autism and Developmental Disorders. 2003; 33(3):239–251. [PubMed: 12908827]

Loveland, K.; Tunali-Kotoski, B. The school-age child with autism. In: Cohen, D.; Volkmar, F., editors. Handbook of Autism and Pervasive Developmental Disorders. Wiley; New York, NY: 1997. p. 283-308.

Luyster RJ, Kadlec MB, Carter A, Tager-Flusberg H. Language assessment and development in toddlers with autism spectrum disorders. Journal of autism and developmental disorders. 2008; 38(8):1426–1438. [PubMed: 18188685]

Mayes L, Volkmar F, Hooks M, Cicchetti D. Differentiating Pervasive Developmental Disorder-Not Otherwise Specified from autism and language disorders. Journal of Autism and Developmental Disorders. 1993; 23:79–90. [PubMed: 7681820]

Muthen, LK.; Muthen, BO. M-plus User's Guide. Muthen and Muthen; Los Angelos, CA: 1998.

Paul R, Orlovski S, Marchinko H, Volkmar F. Conversational behaviors in youth with high-functioning autism and asperger syndrome. Journal of Autism and Developmental Disorders. 2009; 39:115–125. [PubMed: 18607708]

Pettit GS, McClaskey CL, Brown MM, Dodge KA. The generalizability of laboratory assessments of children's socially competent behavior in specific situations. Behavioral Assessments. 1987; 9:81–96.

Rescorla L, Roberts J, Dahlsgaard K. Late talkers at 2: Outcome at age 3. Journal of Speech and Hearing Research. 1997; 40:556–566.

Reynell, J.; Gruber, C. Reynell Developmental Language Scales. Western Psychological Services; Los Angeles: 1990.

Roberts J, Rice M, Tager-Flusberg H. Tense marking in children with autism. Applied Psycholinguistics. 2004; 25:429–448.

Rogers, SJ. Evidence-based practices for language development in young children with autism. In: Charman, T.; Stone, W., editors. Social and Communication Development in Autism Spectrum Disorders. Guilford; New York: 2005. p. 143-179.

Rutter, M.; Le Couteur, A.; Lord, C.; MacDonald, H.; Rios, P.; Folstein, S. Diagnosis and subclassification of autism: Concepts and instrument development. In: Schopler, E.; Mesibov, GB., editors. Diagnostic and Assessment Issues in Autism. Plenum Press; New York: 1988. p. 239-260.

Scarborough HS. Index of productive syntax. Applied Psycholinguistics. 1990; 11:1–22.

Semel, E.; Wiig, EH.; Secord, WA. Clinical Evaluation of language fundamentals-Fourth Edition (CELF-4). The Psychological Corporation; San Antonio: 2003.

Smith T, Groen A, Wynn J. Randomized trial of intensive early intervention for children with pervasive developmental disorder. American Journal of Mental Retardation. 2000; 105(4):269–85. [PubMed: 10934569]

Spence SJ, Cantor RM, Chung L, Kim S, Geschwind DH, Alarco'n G. Stratification Based on Language-Related Endophenotypes in Autism: Attempt to Replicate Reported Linkage. American Journal of Medical Genetics Part B (Neuropsychiatric Genetics). 2006; 141B:591–598.

Tager-Flusberg H. "Once upon a ribbit": Stories narrated by autistic children. British Journal of Developmental Psychology. 1995; 13:45–59.

Tager-Flusberg, H. Language impairments in children with complex neurodevelopmental disorders: The case of autism. In: Levy, Y.; Schaeffer, J., editors. Language competence across populations: Toward a definition of specific language impairment. Lawrence Erlbaum; Mahwah, NJ: 2003. p. 297-321.

Tager-Flusberg H, Rogers S, Cooper J, Landa R, Lord C, Paul R, et al. Defining spoken language benchmarks and selecting measures of expressive language development for young children with autism spectrum disorders. Journal of Speech, Language, and Hearing Research. 2009; 52(3):643–652.

Tager-Flusberg H, Skwerer DP, Joseph RM. Model syndromes for investigating social cognitive and affective neuroscience: a comparison of Autism and Williams syndrome. Social Cognitive and Affective Neuroscience. 2006; 1(3):175–82. [PubMed: 18985104]

Tomblin, JB.; Zhang, X.; Weiss, A.; Catts, H.; Weismer, SE. Dimensions of individual differences in communication skills among primary grade children. In: Rice, ML.; Warren, SF., editors. Developmental Language Disorders: From Phenotypes to Etiologies. Lawrence Erlbaum Associates; Mahwah, NJ: 2004. p. 53-76.

Venter A, Lord C, Schopler E. A follow-up study of high-functioning autistic children. Journal of Child Psychology and Psychiatry. 1992; 33(3):489–507. [PubMed: 1577895]

Ward, A.; Stoker, HW.; Murray-Ward, M. Educational Measurement: Theories and Applications. University Press of America, Inc.; Lanham, MA, US: 1996.

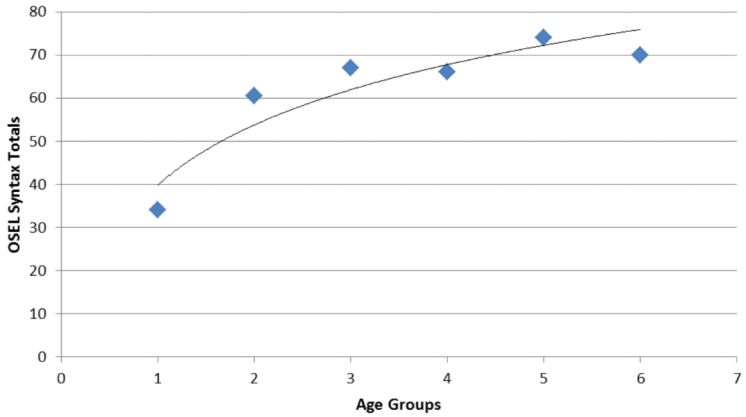

Figure 1.

Fitting a Smooth Line to Derive Age Equivalents for the OSEL Syntax Totals for Males. Note. Age equivalent scores were calculated based on the smooth line (y= 22.777In(x) + 30.189). The fit of this line for the data was $R^2 = 0.857$.

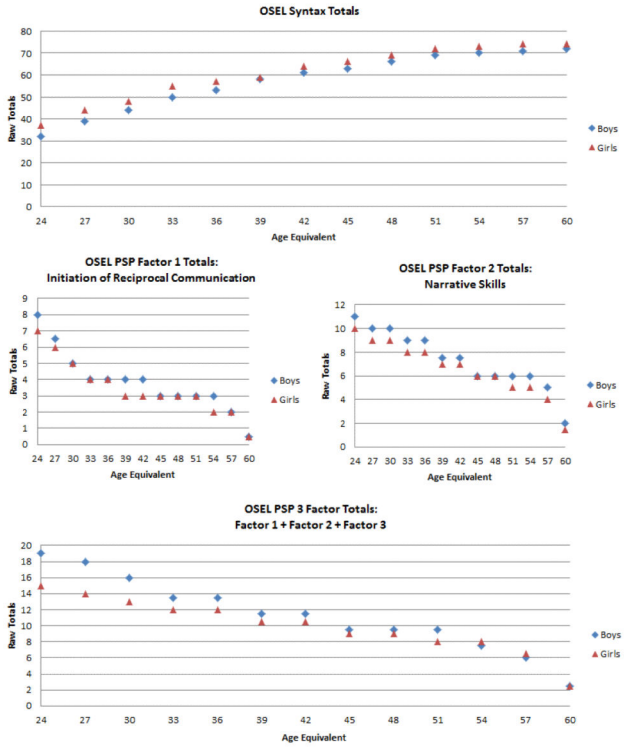

Figure 2.
Age Equivalents (Months) Corresponding to OSEL Syntax and PSP Totals

Table 1

OSEL Tasks

Tasks
1. Play Figure Assembly (“The Bowler Family”)
2. Clarifications*
3. Telling a Picture Story
4. Conversation*
5. Camping Trip/Picnic
6. “Where is it?” Game
7. Retell a Story: Where Are My French Fries?
8. Picture Description

*Note. Clarifications and Conversation tasks occur throughout the administration

Table 2

Participant Characteristics by Age Groups and Gender

	Age Group 1		Age Group 2		Age Group 3		Age Group 4		Age Group 5		Age Group 6		Age Group 7	
Age Range*	24-27		28-30		31-35		36-41		42-47		48-53		54-60	
Gender	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
N	11	10	12	11	12	13	18	12	11	12	17	14	14	13
Mean (SD)														
Age*	25.5 (1.1)	25.3 (1.2)	29.1 (0.8)	28.9 (0.8)	34 (1.6)	34.2 (1.7)	39.5 (1.9)	38.8 (1.7)	45.6 (1.8)	44.6 (2)	51.5 (1.9)	51.2 (1.8)	57.7 (1.8)	57.5 (1.7)
SBIS ABVIQ (average = 10)	11 (1.9)	11.4 (2.1)	9.5 (3.1)	10.8 (3.3)	10.3 (1.7)	11 (2.8)	10.6 (3.5)	11.25 (2.9)	9.4 (1.8)	9.9 (1.8)	11.2 (3.3)	9.6 (2.7)	9.7 (3.1)	9.1 (2.3)
SBIS ABNVIQ (average = 10)	9.34 (2.9)	10.7 (4.7)	8 (4.3)	8.5 (3.9)	10.92 (3.3)	11.5 (2.6)	10.9 (2.9)	11 (2.8)	10 (2.5)	9.5 (2)	9.5 (2.2)	11.6 (4.4)	10.5 (2.5)	12.4 (3.8)
PLS Expressive Communication	109 (19.6)	116.9 (21.5)	110.8 (20.5)	123.1 (20.1)	119.1 (12.5)	117.2 (12.9)	113.2 (19.7)	127.5 (16.5)	108.6 (8.9)	115.6 (15.7)	109.1 (11.6)	113.5 (13.7)	107.8 (10.2)	113.9 (19)
PLS Auditory Comprehension	111.4 (10.8)	109.5 (15)	107.3 (19)	119.5 (13.4)	116.7 (14.5)	115.3 (10.3)	108.8 (14.7)	120.4 (11.6)	110 (12)	114.3 (13.8)	107.2 (11.8)	113.7 (11.4)	114.6 (9.6)	114.1 (16)
CASL Syntax Construction#							103.5 (15.4)	101 (12.5)	96.8 (11.4)	100 (16.3)	106.6 (15.1)	104.6 (15.2)	102.9 (13.2)	106 (21.6)
CASL Pragmatic Judgment#							99.1 (14.3)	100.5 (11.1)	99.9 (15.9)	109.4 (14.7)	109.7 (16.3)	114.1 (14.8)	107.3 (12.3)	116 (24.3)
VABS Expressive Language AE#	32.36 (7.3)	39 (12.9)	36.42 (5.9)	36.4 (10.3)	50.8 (10.2)	41.7 (12.3)	54.6 (12.7)	53.8 (6.7)	52.2 (10.1)	61.8 (8.5)	64.6 (12.5)	66.1 (11.9)	62 (18.6)	67.9 (16.3)

*Note. Age in months;

#CASL is only administered to children from 36 months, and out of 180 children, 5 children whose Vineland data was not available were excluded; *SBIS ABVIQ* Stanford Binet Intelligence Scale Abbreviated Battery Verbal IQ (Verbal Knowledge), *SBIS ABNVIQ* Stanford Binet Intelligence Scale Abbreviated Battery nonverbal IQ (Nonverbal Fluid Reasoning), *PLS* Preschool Language Scale, CASL Comprehensive Assessment of Spoken Language, *VABS* Vineland Adaptive Behavior Scales, *AE* Age Equivalents.

Table 3

Factor Structure of the OSEL Syntax Items

	Factor Loadings
	EFA
Articles	0.83
Regular Plurals	0.75
Irregular Plurals	0.52
Regular Past Verbs	0.79
Irregular Past Verbs	0.9
Progressive Tense Verbs	0.72
Future Tense Verbs	0.9
Copula Verbs	0.82
Modal Auxiliary Verbs	0.87
Infinitive Phrases	0.85
Negation	0.85
Prepositions	0.66
Longest Sentence (number of words)	0.95
Subjective Pronouns	0.89
Objective Pronouns	0.77
Possessive Pronouns	0.76
Subordination	0.9
Coordination	0.93
Adjectives	0.78
Nouns	0.8
Verbs	0.94
Responses to WH- questions	0.55
Responses to Y/N questions	0.67
Questions	0.66
	CFI
	0.99
	RMSEA
	0.057

Note. *EFA* Exploratory Factor Analysis, *CFI* Comparative Fit Index, *RMSEA* Root Mean Square Error Approximation.

Table 4

Factor Structure of the OSEL Pragmatic Semantic Profile Items

	Factor Loadings	
	EFA	**CFA**
Factor 1: Initiation of Reciprocal Communication		
Verbal requests to get needs met	0.64	0.54
Asks for information about thoughts, feelings, or experiences	0.26	0.71
Comments or offers information about thoughts, feelings, or experiences	0.37	0.88
Maintains a conversation	0.59	0.98
(Absence of) Preoccupation with specific interests	0.43	0.87
Factor 2: Narrative Skills		
Repairs/Request clarification	0.19	0.71
Reports main ideas	0.93	0.9
Reports sequence of events/story	0.93	0.96
Comments on characters' emotional and/or mental states	0.31	0.6
Synthesizes cause-and-effect information	0.61	0.93
Factor 3: Unusual Features		
Interrupts the examiner or dominates conversations	0.63	0.39
Stereotyped/Idiosyncratic use of words or phrases	N/A	0.99
Unspecific language and/or semantic errors	0.94	0.88
Immediate echolalia	0.89	0.75
Impolite or inappropriate language	0.37	0.11
	CFI	
	0.995	0.978
	RMSEA	
	0.048	0.080

Note. *EFA* Exploratory Factor Analysis, *CFA* Confirmatory Factor Analysis, *CFI* Comparative Fit Index, *RMSEA* Root Mean Square Error Approximation.

Table 5

OSEL Score Distribution by Age Groups and Gender

	Age Group 1		Age Group 2		Age Group 3		Age Group 4		Age Group 5		Age Group 6		Age Group 7	
Age Range in Months	24-27		28-30		31-35		36-41		42-47		48-53		54-60	
Gender	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
Number of Children	11	10	12	11	12	13	18	12	11	12	17	14	15	13
Mean (Standard Deviation)														
OSEL Syntax	30.1 (12.1)	38.4 (19.9)	31.5 (10.8)	42.4 (13.8)	60.3 (11.9)	58.9 (15)	63.3 (16.7)	64.7 (17.7)	66 (11.2)	69.3 (12.3)	68.2 (17.7)	77 (7.6)	72.3 (9.6)	73.5 (11.3)
PSP Factor 1: Initiation of Reciprocal Communication	6.7 (1.7)	5.5 (2)	7.2 (0.8)	5 (1.6)	3.1 (1.9)	3.8 (1.8)	3.6 (2.1)	3.2 (2.2)	3.3 (1.6)	3.2 (2.7)	2.8 (2.5)	1.9 (1.3)	2.5 (2.1)	2 (1.1)
PSP Factor 2: Narrative Skills	9.4 (1.8)	9.1 (1)	9 (1.5)	9.1 (1.2)	7.4 (2.3)	7.7 (2.2)	6.8 (2.8)	6.3 (2.4)	6.2 (2.4)	6.4 (2.7)	5.8 (2.6)	4.9 (1.9)	4.3 (2.2)	3.8 (2.5)
PSP Factor 3: Unusual Features	1.3 (1.4)	1.5 (1.6)	2.3 (2.3)	2 (3.1)	1 (1.5)	0.7 (1)	0.1 (0.3)	0.1 (0.3)	0.5 (0.5)	0.4 (0.7)	0.5 (0.8)	0.7 (0.7)	0.1 (0.5)	0.4 (0.7)
PSP 3 Factor Total	17.4 (3.4)	16.1 (3.5)	18.4 (2.6)	16.1 (5)	11.5 (5)	12.2 (3.7)	11.1 (4.5)	9.6 (4.2)	9.9 (3)	10 (4.7)	9 (4.7)	7.5 (2.9)	6.9 (3.5)	6.2 (2.9)