RESEARCH ARTICLE

Uh and *um* in Children With Autism Spectrum Disorders or Language Impairment

Kyle Gorman, Lindsay Olson, Alison Presmanes Hill, Rebecca Lunsford, Peter A. Heeman, and Jan P. H. van Santen

Atypical pragmatic language is often present in individuals with autism spectrum disorders (ASD), along with delays or deficits in structural language. This study investigated the use of the "fillers" *uh* and *um* by children ages 4–8 during the autism diagnostic observation schedule. Fillers reflect speakers' difficulties with planning and delivering speech, but they also serve communicative purposes, such as negotiating control of the floor or conveying uncertainty. We hypothesized that children with ASD would use different patterns of fillers compared to peers with typical development or with specific language impairment (SLI), reflecting differences in social ability and communicative intent. Regression analyses revealed that children in the ASD group were much less likely to use *um* than children in the other two groups. Filler use is an easy-to-quantify feature of behavior that, in concert with other observations, may help to distinguish ASD from SLI. ***Autism Res** 2016, 00: 000–000.* © 2016 International Society for Autism Research, Wiley Periodicals, Inc.

Keywords: autism spectrum disorders; language impairment; social communication; conversational reciprocity; pragmatic language; disfluency; fillers

Introduction

Language abilities in children with autism spectrum disorders (ASD) are highly variable [Tager-Flusberg and Joseph, 2003; Whitehouse et al., 2008] although delays and deficits are relatively common [Leyfer et al., 2008; Loucas et al., 2008]. Recent studies suggest that a majority of verbally fluent children with ASD have impairments in *structural language*, which includes phonology, vocabulary, and grammar. In contrast, *pragmatic languag*—the socially-oriented elements of language use—is thought to be universally impaired in ASD [Kim et al., 2014; Klin et al., 2005; Landa, 2000; Lord and Paul, 1997; Tager-Flusberg et al., 2005; Volden et al., 2009]. Yet, there is little consensus on how pragmatic language abilities should be defined or quantified [Russell and Grizzle, 2008; Volden and Phillips, 2010]. More generally, speech communication is critical for everyday functioning, so there is great potential value for interventions that might increase the capacity of an individual with ASD to understand and be understood [Klin et al., 2007].

In this study, we investigated one quantifiable feature of pragmatic language: the use of *uh* and *um*. These "fillers" (or "filled pauses") are subtle yet ubiquitous features of spontaneous speech, accounting for approximately one percent of word tokens produced by typical adults [Acton 2011; see also Fox Tree, 1995, p. 709]. Like other types of disfluency (including pauses, false starts, repetitions, and repairs), fillers are thought to reflect difficulties with planning and delivering speech [Clark, 1994; Levelt, 1989, p. 484]. For instance, fillers are particularly common immediately before pauses in speech [Clark and Fox Tree, 2002]. But there is extensive evidence that fillers also act as interpersonal displays directed at the listener [Sacks, Schegloff, & Jefferson, 1974]. On hearing an *uh* or *um*, listeners may infer that the speaker is experiencing difficulty with word retrieval or speech planning [Maclay and Osgood, 1959; Stenström, 1994, p. 76f.], and will often provide verbal assistance to the speaker [Clark and Wilkes-Gibbs, 1986; Jefferson, 1974]. Listeners may also use fillers to uncover linguistic structure during speech perception. For instance, fillers may be used as cues to major syntactic boundaries [Bailey and Ferreira, 2003; Martin and Strange, 1968; Swerts, 1998] or to new information being introduced into the discourse [Arnold et al.,

From the Oregon Health & Science University, 3181 SW Sam Jackson Park Rd., GH40, Portland, Oregon 97239-3098 (K.G.,); Center for Spoken Language Understanding, Institute on Development and Disability, Oregon Health & Science University, Portland, Oregon (K.G., L.O., A.P.H., R.L., P.A.H., J.P.H.v.S.,); Department of Pediatrics, Oregon Health & Science University, Portland, Oregon (K.G., A.P.H., P.A.H., J.P.H.v.S.,); BioSpeech Inc., Portland, Oregon (L.O., R.L., P.A.H., J.P.H.v.S.)
Received November 18, 2014; accepted for publication October 05, 2015
Address for correspondence and reprints: Kyle Gorman, Oregon Health & Science University, 3181 SW Sam Jackson Park Rd., GH40, Portland, OR 97239-3098. E-mail: kylebgorman@gmail.com
Published online 00 Month 2016 in Wiley Online Library (wileyonlinelibrary.com)
DOI: 10.1002/aur.1578

2003; Kidd et al., 2011]. Listeners also may use fillers as cues to the speaker's mental state; for example, high rates of filler use often causes listeners to infer that the speaker lacks knowledge about the topic under discussion [Brennan and Clark, 1995; Smith and Clark, 1993]. Thus, regardless of the speaker's intent, fillers can influence how a listener perceives and responds to a speaker during conversation.

In summary, filler use appears to be an important component of conversational reciprocity, which is thought to be impaired in ASD [e.g., Tager-Flusberg et al., 2005]; for example, compared to typically developing (TD) children, children with ASD have difficulty initiating conversation [Tager-Flusberg, 1996], responding appropriately to the initiations of others [Adams et al., 2002; Capps et al., 1998; Stone and Caro-Martinez, 1990], taking turns [Botting and Conti-Ramsden, 2003; Ramberg et al., 1996], and staying on topic [Capps et al., 1998; Lam and Yeung, 2012; Losh and Capps, 2003; Loveland et al., 1990; Paul et al., 2009]. We thus hypothesized that difficulties with conversational reciprocity would also be reflected in atypical filler use in children with ASD.

Other social-cognitive impairments in ASD may also contribute to atypical filler use, for instance executive functioning difficulties, which are common in individuals with ASD [Kenworthy, Yerys, Anthony, & Wallace, 2008]. One executive function (inhibition) has been linked to disfluency in typical adults [Engelhardt et al., 2013] and other studies of typical adults have found that cognitive load increases disfluency rates [e.g., Bortfeld et al., 2001; Christenfeld, 1995; Engelhardt et al., 2010; Schachter et al., 1991]; presumably, cognitive load attenuates speakers' ability to monitor their speech planning for anticipated delays [Bock, 1982; Levelt, 1983]. Adams et al. [2002] suggests that "social-emotional conversation" may be particularly loading for children with ASD. In that study, children with Asperger syndrome produced more "pragmatically problematic" responses (roughly, those judged to be pragmatically inappropriate in context) than peers with conduct disorder, but only during social-emotional conversations. Taken together, these studies suggest that filler use in ASD may vary as a function of the social demands of the topic under discussion.

A number of studies have attempted to confirm a clinical impression that high-functioning children with ASD "may lack in fluency" [Klin et al., 2005, p. 99] compared to TD children. Thurber and Tager-Flusberg [1993] examined disfluency in ten children and adolescents with autism during narration of a wordless picture book. Participants with autism produced fewer within-phrase pauses than participants with typical development or mild mental retardation, but no group differences in repetitions or false starts were found. In another study, Lake et al. [2011] elicited conversational speech from 13 adults with ASD. Compared to age-matched controls, ASD participants produced more disfluent repetitions and more pauses, but fewer revisions and fillers. Suh et al. [2014] used a storytelling task to elicit speech from children with high-functioning ASD, children with a past diagnosis of ASD who no longer met criteria for ASD [see Suh et al., 2014 for details], and TD children. They found that both clinical groups produced higher rates of repetitions and revisions than TD children, but filler rates were comparable across groups. These conflicting results are likely due to methodological differences. For example, Suh et al. studied children using a storytelling task to elicit speech, whereas Lake et al. elicited speech by asking their adult participants about their interests and hobbies. In addition, neither study controlled for participants' language abilities. Indeed, Lake et al. found that one measure of language ability—mean length of utterance—was correlated with disfluency rate, and similar results have been found in studies of typical adults [Bortfeld et al., 2001; Cook, Smith, & Lalljee, 1974; Oviatt, 1995; Shriberg. 1996] and children with specific language impairment (SLI) [Thordardottir & Weismer, 2002]. Thus, group differences may have been due to the lower average language abilities of the ASD group rather than to any specific feature of ASD. In summary, it is still unclear whether children with ASD use fillers differently than TD children or children with language delays not related to ASD.

Although the above studies conflated the fillers *uh* and *um*, Clark and Fox Tree [2002] argue that they have different functions: *uh* serves to signal minor delays, whereas *um* signals major delays. The evidence for this is primarily distributional. In typical adults, *um* is more often followed by a pause than *uh*, and when a pause is present after a filler, it tends to be longer after *um* than after *uh*. Children with ASD also exhibit this pattern, producing more pauses, and longer pauses, after *um* than after *uh* [de Villiers, 2011; Lunsford, Heeman, Black, & van Santen, 2010]. Clark and Fox Tree also report that *um* is the more common of the two fillers at the start of intonational phrases (e.g., *um, we went to the beach*)—where speech planning demands are presumably at their greatest—whereas *uh* is more common elsewhere (*we went to the, uh, beach*). In summary, *uh* and *um* have different usage patterns, and perhaps, different functions in discourse.

The Present Study

Heeman et al. [2010], Lunsford et al. [2010], and Lunsford [2012] hypothesized that children with ASD would produce atypical patterns of filler use compared to TD children, and preliminary analyses conducted on a smaller subgroup of children included in the current sample provided initial support for this hypothesis. However, these prior analyses did not address specific hypotheses regarding the effects of topic on filler use, and did not examine associations between filler use and individual differences in cognitive abilities or ASD symptoms.

In the present study, we re-examined this hypothesis using a larger sample of children with and without ASD, according to best estimate clinical (BEC) diagnoses. For comparison, we included a TD group as well as a group of children with SLI, a neurodevelopmental disorder defined by language delays or deficits in the absence of other developmental or sensory impairments [Tomblin, 2011]. SLI is associated with deficits in structural language whereas ASD involves atypicalities in both structural and pragmatic language [Shulman and Guberman, 2007]. To determine whether there is a specific ASD-related profile for filler use, the SLI clinical group is essential [Bishop, 2001; Ellis Weismer, 2013; Kjelgaard and Tager-Flusberg, 2001]; otherwise, observed group differences may be attributed to difficulties with language that are common in, but not specific to, children with ASD.

Methods

Participants

One hundred and ten children from the Portland, OR metropolitan area, between 4 and 8 years of age, took part in the study: 50 children with ASD (45 male), 43 TD children (TD; 31 male), and 17 children with SLI (11 male).

Recruitment and screening. Participants with ASD were recruited through local healthcare specialists, educational service districts, autism clinics, parent groups, and local nonprofit autism organizations. Participants with SLI were recruited through local speech clinics, speech-language pathologists, and the Oregon Speech and Hearing Association. Advertisements were also placed in local newspapers and on "community tables" at local elementary schools. All participants had full-scale IQ scores of 70 or higher on the Wechsler Preschool and Primary Scale of Intelligence [WPPSI-III; Wechsler, 2002] for children under 7 years of age, or the Wechsler Intelligence Scale for Children [WISC-IV; Wechsler, 2004] for children ages 7 or older. Children were excluded if they had any of the following: (a) known metabolic, neurological, or genetic disorder, (b) gross sensory or motor impairment, (c) brain lesion, (d) orofacial abnormalities (such as cleft palate), or (e) mental retardation. All participants spoke English as their first language, and produced a mean length of utterance in morphemes (MLUM) of at least three. During the initial screening, a certified speech-language pathologist confirmed the absence of speech intelligibility impairments.

Diagnostic groups. BEC judgment by experienced clinicians is thought to be the gold standard for ASD diagnosis [e.g., Klin et al., 2000; Spitzer and Siegel, 1990]. In this study, a panel of clinicians, including two clinical psychologists, a speech-language pathologist, and an occupational therapist, all of whom had clinical expertise with ASD, based their judgments on DSM-IV-TR criteria [American Psychiatric Association, 2000] for ASD. Only children who received a consensus BEC diagnosis of ASD were included in this study. The consensus diagnosis was confirmed by above-threshold scores on the Autism Diagnostic Observation Schedule-Generic [ADOS-G; Lord et al., 2000] according to the revised algorithms [Gotham et al., 2007] and the Social Communication Questionnaire [SCQ; Rutter et al., 2003] according to the cutoff score of 12 recommended for research purposes [Lee et al., 2007]. There were a small number of nonresponses on questions on the SCQ (corresponding to 0.3% of the overall data). Before computing SCQ scores, chained equation multiple imputation [Su et al., 2011] was used to fill in nonresponses (corresponding to 0.3% of the overall data).

Language impairment was assessed using the Clinical Evaluation of Language Fundamentals (CELF), a test which produces a composite summary of expressive and receptive language abilities. For children younger than 6 years of age, the CELF Preschool-2 [Semel, Wiig, & Secord, 2004] was administered; the CELF-4 [Semel et al., 2003] was used for children age 6 or older. Language impairment was determined by a CELF core language score (CLS) more than one standard deviation below the mean. Half of the 50 children with ASD were identified as language impaired according to this criterion. Children assigned to the SLI group also were required to have (a) a documented history of language delays or deficits, and (b) a BEC consensus judgment of language impairment but not ASD, taking into account medical and family history, assessments performed as part of this study or at an earlier time by others, and school records. Children with a BEC diagnosis of SLI were excluded from the study if they reached threshold on both the ADOS-G and the SCQ.

Children who did not meet the criteria for either ASD or SLI were assigned to the TD group, but were excluded from the study if they had any family members diagnosed with either ASD or SLI, a history of psychiatric disturbance (e.g., ADHD) or if the child was above threshold according to the ADOS-G or the SCQ.

Procedures

Participants completed a battery of experimental tasks and cognitive, language, and neuropsychological assessments over six sessions of 2–3 hr each. All procedures were approved by the Oregon Health & Science

University Institutional Review Board. Participating families were fully informed about the study procedures and provided written consent.

Standardized measures. The ADOS [Lord et al., 2000], a semistructured autism diagnostic observation, was administered to all children in the current study, and was scored according to the revised algorithms [Gotham et al., 2007]. Ten children received ADOS Module 2, and 100 received Module 3.[1] Domain calibrated severity scores (CSS) were calculated as indications of severity of social affect (SA) and restricted and repetitive behavioral (RRB) symptoms (Hus et al., in press). The social affect calibrated severity scores (ADOS SA; range: 1–10) was used as a clinician-reported measure of social communication difficulties. Transcripts of the ADOS were used to derive several other measures (see next section).

Verbal IQ (VIQ), performance IQ (PIQ), and full-scale IQ (FSIQ) were estimated using the Wechsler scales tests, as described above.

Parents completed the behavior rating inventory of executive function [BRIEF; Gioia et al., 2000] for children 6 years of age or older, and the BRIEF-Preschool Version [BRIEF-P; Gioia et al., 2003] for children under 6. Both forms were used to compute the global executive composite (GEC), which was used as a measure of overall executive functioning.

Structural language abilities in children with ASD and SLI were assessed using the CELF CLS, as well as the two CELF subscales, the expressive language index (ELI) and the receptive language index (RLI). TD children were screened for language impairment but did not complete the CELF.

Parents completed the Children's Communication Checklist [CCC-2; Bishop, 2003], a 70-item questionnaire assessing the child's communication abilities in natural settings. The general communication composite (GCC) is the sum of subscale scores from the eight CCC-2 domains related to communication (speech, syntax, semantics, coherence, initiation, scripted language, context, and nonverbal communication). The social-interaction deviance index (SIDI) uses these subscales to measure relative strengths in structural vs. pragmatic language; a negative SIDI indicates stronger structural language abilities, and a positive score indicates stronger pragmatic language abilities.

[1]An anonymous reviewer asks whether data gathered from separate ADOS modules can be compared. To test whether our results were sensitive to this, we repeated all statistical analyses excluding data from children who completed the ADOS Module 2. The results indicated that excluding data from children who completed Module 2 had no effect on the results obtained.

Parents also completed the Social Communication Questionnaire [Rutter et al., 2003], a 40-item assessment of symptomatology associated with ASD. The SCQ communication total score [SCQ-CTS; range: 0–12; Berument et al., 1999], the sum of scores for items in the communication domain, was used as an additional parent-reported measure of communication abilities.

Transcription. ADOS sessions were recorded and the child and examiner's speech was transcribed using Praat software. Transcribers were blind to study hypotheses and to participants' diagnostic status and intellectual abilities. The transcriptions were generated using a subset of the systematic analysis of language transcripts (SALT) guidelines [Miller and Chapman, 1985]. Transcribers were instructed to mark mazes (i.e., disfluent intervals of speech), including sequences of fillers and false starts, repetitions, and revisions. Each ADOS transcription was segmented into four "activities": Play (including Make-Believe Play and Joint Interactive Play), Description of a Picture, Telling a Story from a Book, and Conversation. For children who received the ADOS Module 3, the Conversation activity included the Emotions, Social Difficulties and Annoyance, Friends and Marriage, and Loneliness sections. The remaining portions of the ADOS were not transcribed. Within each activity, conversational turns were segmented into individual utterances (or "C-units"), each consisting of (at most) a main clause and any subordinate clauses modifying it.

Measures Derived from ADOS Transcripts. ADOS transcripts were used to compute overall MLUM [Brown 1973] using SALT software [Miller and Chapman, 1985]. MLUM is a simple, face-valid measure of morphological and syntactic complexity recommended as a benchmark of spoken language development in children with autism [e.g., Tager-Flusberg et al., 2009]. These transcripts were also used to count *uh*s, *um*s, and fluent words (words which are not part of a maze) for each participant. In the case that a child produced multiple consecutive fillers within a single utterance (e.g., *she had the um um starfish*), only the first filler was counted. Immediately-repeated fillers were excluded on the hypothesis that a repeated filler is not statistically independent of the preceding filler, and thus their inclusion would violate the independence assumptions of the quantitative analyses.[2] The full data set contained 1,261 tokens of *uh* and 2,523 tokens of *um*.

[2]An anonymous reviewer asks whether the exclusion of immediately-repeated fillers may have influenced the results. To test this, we repeated all statistical analyses without these exclusions and found that it had no effect on the results obtained.

Table 1. Interannotator Agreement

	Annotator 1		Annotator 2	
	Accuracy	κ	Accuracy	κ
Filler type	0.949	0.893	0.910	0.810
Filler presence/absence	0.871	0.737	0.865	0.729

Interannotator agreement accuracy and Cohen's kappa (κ) for filler type (*uh* or *um*), and filler presence/absence.

Filler annotation quality was assessed retrospectively using a stratified random sample of the full dataset. The sample contained four utterances per child, two of which had been transcribed as containing one or more fillers (*uh* or *um*) and two of which had been transcribed without any fillers. Utterances were excluded from this sample if they contained unintelligible words or if the examiner's speech overlapped the child's speech. Audio files of these sample utterances were extracted with a two-second fade-in/fade-out. These files were then transcribed, according to the same guidelines, by two experienced transcribers, neither of whom participated in the initial transcription efforts. Both transcribers were blind to participant identity and group assignment.

When original and retrospective annotations both contained a filler, the original and retrospective annotators transcribed the same filler type (*uh* or *um*) in 95% and 91% of cases, with Cohen's kappa (κ) of 0.893 and 0.810, respectively (see Table 1). This corresponds to "almost perfect" agreement according to Landis and Koch's [1977, p. 165] qualitative guidelines. These retrospective transcriptions were also used to assess agreement for the presence or absence of fillers, ignoring filler type. In 87% of cases, the original and retrospective transcriptions agreed on presence or absence of fillers, with κ of 0.737 and 0.729, respectively, indicating "substantial" agreement according to the Landis and Koch guidelines.

Statistical Analysis

Inferential analyses were conducted using mixed effects logistic regression [Pinheiro and Bates, 2000] with a per-subject random intercept. Compared to conventional (i.e., fixed effects) logistic regression, this method provides a principled solution to problems of non-independence and heteroscedascity arising when subjects contribute different numbers of observations, as is the case here. The primary independent variable was participant group (ASD, SLI, or TD). All models also included three participant-linked covariates: chronological age, full-scale IQ, and MLUM. Each token was coded for ADOS activity (Play, Description of a Picture, Telling a Story from a Book, or Conversation). A binary predictor was used to code whether a token was utterance-initial or noninitial. To facilitate interpretation, continuous variables were z-transformed, and sum coding was used to encode categorical variables. The log-likelihood ratio test was used to test for significance of individual predictors, and the Tukey HSD test was used to test for significant differences within factor groups. Exploratory analyses were conducted by measuring correlations with Kendall's τ_b, a nonparametric correlation statistic.

Results

Group Characteristics

Summary statistics for the three diagnostic groups are reported in Table 2.

Inferential Analyses

Three separate inferential analyses were performed. The first two compared children's productions of *uh* and *um*, respectively, to their productions of fluent words,

Table 2. Group Statistics for the Sample

	ASD ($n = 50$)		SLI ($n = 17$)		TD ($n = 47$)			
	mean	(s.d.)	mean	(s.d.)	mean	(s.d.)	$P(F)$	$P(\text{HSD}) < 0.05$
CA	6.6	(1.2)	7.1	(1.1)	6.2	(1.2)	0.038	TD < SLI
FSIQ	98.3	(15.8)	88.3	(8.0)	119.3	(11.7)	<0.001	SLI < ASD < TD
VIQ	95.1	(17.8)	85.8	(6.2)	119.3	(12.9)	<0.001	SLI = ASD < TD
PIQ	108.6	(17.2)	101.6	(11.3)	118.2	(14.6)	<0.001	SLI = ASD < TD
CLS	89.7	(21.6)	74.2	(8.4)	n.a.	(n.a.)	0.006	SLI < ASD
MLUM	4.2	(1.0)	4.1	(1.0)	4.9	(0.9)	<0.001	SLI = ASD < TD
GCC	50.6	(10.8)	47.6	(12.5)	96.1	(13.8)	<0.001	SLI = ASD < TD
GEC	69.2	(9.0)	65.7	(13.4)	44.6	(8.1)	<0.001	TD < SLI = ASD
SCQ	19.6	(4.9)	11.3	(6.7)	2.9	(2.5)	<0.001	TD < SLI < ASD
ADOS	7.5	(2.0)	2.9	(2.8)	1.2	(0.5)	<0.001	TD < SLI < ASD

Mean and standard deviation for each group, P-value for one-way ANOVA on group, and post hoc group contrasts which are significant at $\alpha = 0.05$. CA, chronological age in years; FSIQ, full-scale IQ; VIQ, verbal IQ; PIQ, performance IQ; CLS, CELF CLS (not available for TD); MLUM, mean length of utterance in morphemes; GCC, CCC-2 general communication composite; GEC, BRIEF global executive composite; SCQ, SCQ total score; ADOS, ADOS-G calibrated severity score.

Table 3. Results for Regression on *uh* Rate

	Log-odds	s.e.		
(Intercept)	−5.462	0.11	χ^2	$P(\chi^2)$
Group:			2.27	0.306
ASD	0.191	0.13		
SLI	−0.208	0.19		
TD	0.017			
CA	0.00	0.09	0.00	0.986
FSIQ	0.089	0.12	0.51	0.474
ADOS Activity:			52.67	<0.001
Play	0.047	0.05		
Description of a picture	0.029	0.06		
Telling a story from a book	−0.382	0.07		
Conversation	0.306			
Context:			975.18	<0.001
Initial	0.915	0.03		
Noninitial		−0.915		

Mixed effects logistic regression on *uh* rate; predictors which are positively associated with *uh* have positive log-odds and predictors which are negatively associated with *uh* have negative log-odds. CA, chronological age in years; FSIQ, full-scale IQ.

Table 4. Results for Regression on *um* Rate

	Log-odds	s.e.		
(intercept)	−4.999	0.12	χ^2	$P(\chi^2)$
Group:			16.13	<0.001
ASD	−0.588	0.16		
SLI	0.042	0.23		
TD	0.546			
CA	0.103	0.11	0.82	0.364
FSIQ	−0.054	0.15	0.13	0.719
ADOS Activity:			216.76	<0.001
Play	−0.215	0.04		
Description of a picture	0.265	0.04		
Telling a story from a book	−0.438	0.06		
Conversation	0.388			
Context:			1630.39	<0.001
Initial	0.848	0.02		
Noninitial		−0.848		

Mixed effects logistic regression on *um* rate; predictors which are positively associated with *um* have positive log-odds and predictors which are negatively associated with *um* have negative log- odds. CA, chronological age in years; FSIQ, full-scale IQ.

i.e., those not part of a maze; these test for group differences in the use of these two fillers while controlling for any group differences in fluent verbal output. The third analysis compares children's productions of *uh* to productions of *um*, controlling for any group difference in overall filler production.

***Uh* rate.** In the first mixed effects regression, each token of *uh* was coded as a "hit" and each fluent word (those not part of a maze) as a "miss." The results are

Table 5. Results for Regression on *uh* vs. *um*

	Log-odds	s.e.		
(intercept)	0.433	0.16	χ^2	$P(\chi^2)$
Group:			16.29	<0.001
ASD	−0.795	0.20		
SLI	0.263	0.29		
TD	0.531			
CA	0.124	0.14	0.75	0.387
FSIQ	0.075	0.19	0.16	0.691
ADOS Activity:			24.32	<0.001
Play	−0.262	0.08		
Description of a picture	0.160	0.08		
Telling a story from a book	−0.126	0.11		
Conversation	0.229			
Context:			1.93	0.164
Initial	0.064	0.05		
Noninitial		−0.064		

Mixed effects logistic regression comparing *uh* and *um* frequencies; predictors which favor *um* have positive log-odds and predictors which favor *uh* have negative log-odds. CA, chronological age in years; FSIQ, full-scale IQ.

shown in Table 3. There was no effect of group, chronological age, or full-scale IQ. However, there was a significant effect of ADOS activity ($\chi^2 = 52.67$, $P < 0.001$), and post hoc tests identified significant contrasts between nearly all pairs of ADOS activities, with the highest rate of *uh* occurring during the conversation activity (Telling a Story from a Book < Description of a Picture = Play < Conversation, all $P < 0.001$). Tokens of *uh* were also more likely in utterance-initial position than in non-initial position ($\chi^2 = 975.18$, $P < 0.001$).

***Um* rate.** The second mixed effects regression investigated *um* use. Each token of *um* was coded as a "hit" and each fluent word as a "miss." The results are shown in Table 4. There was a main effect of group ($\chi^2 = 16.13$, $P < 0.001$). Post hoc tests revealed that the ASD group used *um* at a significantly lower rate than children with typical development ($P < 0.001$); differences between the other pairs of groups were nonsignificant. Chronological age, full-scale IQ, and MLUM were not associated with use of *um*. Once again, there were significant effects of ADOS activity ($\chi^2 = 216.76$, $P < 0.001$) and post hoc tests identified significant contrasts between several pairs of ADOS activities (Telling a Story from a Book < Play = Description of a Picture < Conversation, all $P < 0.001$). Tokens of *um* were more likely in utterance-initial position than in non-initial position ($\chi^2 = 1630.39$, $P < 0.001$).

***Uh* vs. *um* use.** The final regression examined filler choice by comparing *uh* and *um* frequencies. Tokens of *um* were coded as "hits" and tokens of *uh* as "misses," thus controlling for any group differences in overall filler

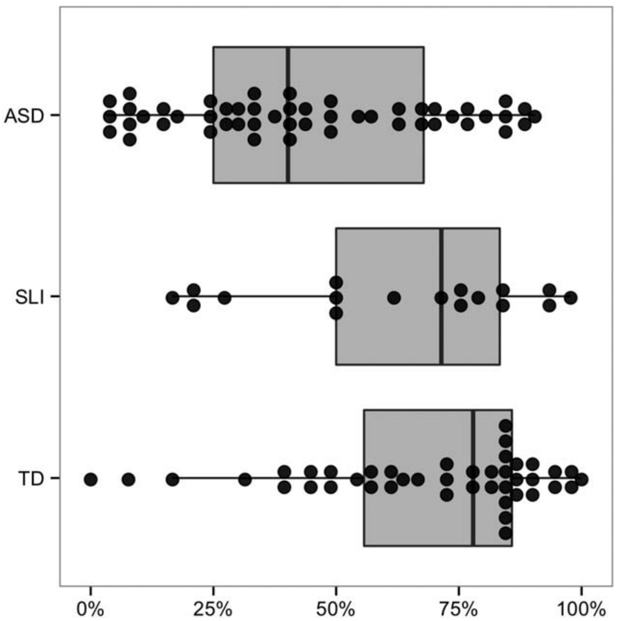

Figure 1. The *x*-axis represents the percentage of fillers which are *um* (rather than *uh*) for each child. The vertical lines indicate boundaries between the group quartiles. Children in the ASD group used fewer *um*s on average than children in the SLI and TD groups.

Table 6. Correlations with *um-uh* Ratio

	ASD	SLI	TD
CA	−0.05	0.13	0.1
FSIQ	0.01	−0.26	0.06
PIQ	0.03	−0.15	0.10
VIQ	−0.01	0.00	−0.01
BRIEF GEC	−0.06	0.01	0.05
MLUM	0.12	0.02	0.34
CELF			
CLS	0.01	−0.29	n.a.
RLI	0.00	−0.08	n.a.
ELI	0.00	−0.25	n.a.
CCC-2			
GCC	0.18	0.23	0.14
SIDI	−0.02	−0.18	0.07
SCQ CTS	−0.29	−0.33	−0.22
ADOS SA	0.04	0.06	−0.19

Associations between per-child *um-uh* ratio and age, intellectual ability, executive function, language, and social ability, as measured by Kendall's τ_b. Children in the TD group did not complete the CELF. CA, chronological age in years; GEC, BRIEF global executive composite; FSIQ, full-scale IQ; VIQ, verbal IQ; PIQ, performance IQ; MLUM, mean length of utterance in morphemes; CLS, CELF core language score; RLI, CELF receptive language index; ELI, CELF expressive language index; GCC, CCC-2 general communication composite; SIDI, CCC-2 social-interaction deviance index; SCQ CTS, SCQ communication total score; ADOS SA, ADOS-G social affect calibrated severity score.

rates. The results are shown in Table 5. There was a main effect of group ($\chi^2 = 16.29$, $P < 0.001$). Post hoc tests revealed that the ASD group used *um* at a significantly lower rate than the TD group ($P < 0.002$) but there were no significant differences between the other pairs of groups. Once again, there was a significant effect of ADOS activity ($\chi^2 = 24.32$, $P < 0.001$). Post hoc tests revealed one significant contrast between activities: compared to Play, Conversation strongly favored *um* ($P <$ 0.001). The *um-uh* comparison is depicted in Figure 1. Each dot represents the percentage of fillers which are *um*—i.e., *um*s/(*uh*s + *um*s)—for each participant.

Associations Between Filler Use and Other Measures

We also conducted an exploratory analysis to investigate whether within-group heterogeneity in filler use was associated with chronological age, intellectual ability, executive function, structural and pragmatic language, or social communication. We computed correlation coefficients for the association between the *um-uh* ratio and each of these measures; separate analyses were conducted for each of the three groups. Within each group, *P*-values were adjusted for false discovery rate [Benjamini and Hochberg, 1995]. Many of these tests are complementary—i.e., several pairs of independent variables are highly correlated and measure closely-related constructs—and the resulting statistical tests are very likely underpowered (particularly in the SLI group), so even the adjusted *P*-values should be interpreted with caution.

The results are shown in Table 6. There were no reliable associations between *um-uh* ratio and chronological age, intelligence, or executive function. There was a significant association between *um-uh* ratio and MLUM in the TD group ($\tau_b = 0.34$, $P = 0.020$). There were no reliable associations between *um-uh* ratio and scores on the CELF, the CCC-2, or the ADOS SA. In all three groups, there were weak negative correlations between *um-uh* ratio and the SCQ communication total score (SCQ CTS), a parent-reported measure. This effect was marginal in ASD ($\tau_b = -0.29$, $P = 0.073$) and nonsignificant in TD and SLI. (Note that higher scores on the SCQ CTS and ADOS SA indicate greater degrees of impairment.)

Discussion

In this study, we investigated the use of *uh* and *um* in a sample consisting of children with ASD, SLI, and typical development. These fillers play a subtle but important role in everyday life and atypical use of fillers by speakers with ASD may contribute to difficulties engaging in conversations with others. Although we did not find any group differences in *uh* rate, we found robust group differences in *um* rate and in *um-uh* ratio. Approximately 40% of the fillers used by children with ASD were *um*, but *um* accounted for more than 70% of the fillers used by children in the TD and SLI groups.

Participants in this study were all highly verbal. However, there was considerable variability in their structural language abilities. Half of the 50 participants with ASD had a CELF CLS more than one standard deviation below the normative mean, as did all participants in the SLI group. On three standard measures of language—MLUM, verbal IQ, and the CCC-2 GCC—the ASD and SLI groups were well-matched, and both groups had significantly lower mean values than the TD group (see Table 2). We found that the SLI and TD groups both had a significantly higher *um-uh* ratio than the ASD group, and were not significantly different from each other. This suggests that a low use of *um* is specific to ASD, but independent of impairments in structural language, which are relatively prevalent among—but not specific to—children with ASD [e.g., Leyfer et al. 2008; Loucas et al. 2008]. The possibility of an autism-specific pragmatic deficit is particularly interesting in light of findings suggesting that pragmatic difficulties are also common in children with SLI [e.g., Bishop, 2001].

We hypothesized that different ADOS activities might influence filler use and choice. In this study, ADOS activity emerged as one of the most robust predictors of filler rates. To take an extreme example, *um* was more than twice as common during the conversation activity than during the Telling a Story from a Book activity, an effect that is nearly as large as differences between the ASD and TD groups (see Table 4). As mentioned earlier, studies of typical adults suggest that cognitive load contributes to elevated disfluency rates. Our results are consistent with this, under the assumption that social-emotional conversation is particularly cognitively loading [cf. Adams et al., 2002], and that the demands of face-to-face discussion of emotional topics limit speakers' abilities to effectively plan and monitor their speech. More generally, this result highlights the importance of controlling for topic in quantitative studies of pragmatic language.

This study had several limitations. First, participants were drawn from a relatively wide age range (4–8). Although chronological age was included as a covariate in regression analyses, developmental differences may have obscured important group differences. Second, the majority of the participants were male. Consequently, we lacked statistical power to investigate gender differences, although there is some evidence that typical adult female speakers produce more *um*s than their male peers [Acton, 2011; Tottie, 2011]. Furthermore, we did not investigate the role of socioeconomic status or ethnicity, although class and ethnicity may play a role in pragmatic language, including use of *uh* and *um* [Rayson, Leech, & Hodges, 1997; Tottie, 2011]. Another limitation is that the diagnostic groups were defined using strict cutoffs for SLI and ASD; different cutoffs might produce different results. Finally, all participants were high-functioning, limiting the generalizability of these results to the larger population of individuals with ASD.

The current study was limited to the English fillers *uh* and *um*. However, the general patterns documented here are not necessarily limited to children acquiring English. Fillers appear to be a linguistic universal [Allwood et al. 1990, p. 33], and virtually all languages have at least two distinct fillers [Clark and Fox Tree, 2002; Wilkins, 1992]. As in English, different fillers tend to exploit different discourse niches. In Dutch, for instance, *uh* [əh] favors phrase-medial position and tends to precede shorter pauses, while *um* [əm] favors phrase-initial position and tends to introduce longer pauses [Swerts, 1998, p. 490], just as in English. Similarly, in Japanese, the fillers *ano* and *sono* pattern with English *uh*, whereas *e* and *eto* pattern with English *um* [Watanabe and Ishi, 2000; Watanabe, 2001, 2002]. Thus, it is possible that similar patterns will be found in children acquiring other languages, although we leave this as a topic for future research.

One potential confound in this study concerns the role that prosody plays in the perception of fillers. In the speech of typical adults, there are distinct prosodic cues associated with *uh* and *um* [Levelt and Cutler, 1983; Shriberg and Lickley, 1993]. Listeners can use these cues to identify fillers, even in unfamiliar languages [Lai et al., 2007] or in speech that has been low-pass filtered [Lickley, Shillcock, & Bard, 1991]. Since many children with ASD are thought to exhibit atypical speech prosody [e.g., Paul et al., 2005; Peppé et al., 2007; van Santen et al., 2010], this might make it more difficult for transcribers—who presumably rely on these prosodic cues when making transcriptions—to detect fillers in these children. It remains to be seen whether disfluencies are associated with different prosodic cues in the speech of children with ASD. However, there were no group differences in overall rates of fillers or annotator reliability, suggesting that atypical prosody in children with ASD cannot fully account for the group differences we found.

It is not known whether individual abilities in production and perception of fillers are associated, but it is possible that in addition to atypical filler use, children with ASD may also have difficulties perceiving fillers (i.e., due to difficulties with prosody perception) or interpreting the full set of social cues that fillers convey. While our data does not directly address speech perception, this may be another fruitful topic for future work.

Our exploratory analysis suggests that *um-uh* ratio is associated with a parent-endorsed measure of social communication ability. Thus, a child's filler use may influence a parent's assessment of his or her child's communicative competence, but further research is needed to investigate parental perceptions of communicative abilities in ASD to determine the extent of this effect.

Our inferential analyses uncovered robust differences between children with and without ASD. Given that social-communicative deficits are a defining feature of ASD, our findings provide convergent evidence for the essentially social function of fillers in the speech of typical individuals [e.g., Brennan and Schober, 2001; Clark, 1994; Fox Tree, 2001]. Our findings also contribute to our understanding of the inherent difference between *uh* and *um*. Similar findings are reported in a recent study by Irvine [2014], who also found that children and adolescents with ASD produce a lower *um* rate than typical controls. In that study, a third group of participants, who received a prior diagnosis of ASD but who no longer meet diagnostic criteria, produced *um* at similar rates to TD peers and at significantly higher rates than peers with ASD. If group differences in *um* use and correlations between *um* use and social communication abilities are replicated, fillers, along with other subtle aspects of pragmatic language, may be a useful target for intervention, particularly in individuals with ASD who are verbal and high-functioning.

Conclusions

We have shown that children with high-functioning ASD produce *um* at a rate much below that of children without ASD, and that use of *um* is positively associated with parents' assessments of their child's social communication abilities. Crucially, use of *um* is not associated with language impairment within the group of children diagnosed with ASD, and children with SLI use *um* at a similar rate as their TD peers. Filler use is an easily quantified feature of pragmatic language that, in concert with other behavioral features, may distinguish ASD and SLI, a notoriously difficult differential diagnosis [Bishop and Norbury, 2002; Bishop et al., 2008; Cox et al., 1999]. We believe that this general approach—using contextualized natural language samples to quantify features of pragmatic language—will provide much-needed conceptual precision and complement existing methods of diagnosis and phenotypic characterization based on clinical observation or structured assessment.

Acknowledgments

We thank Mabel Rice for helpful discussion on criteria for classifying specific language impairment, and Lauren Kenworthy for assistance with measures of executive function. Thanks also to Julianne Myers, Hilary Prichard, Emily Prud'hommeaux, and audiences at IMFAR 2014 for their comments on this work. We also thank the two anonymous reviewers. L. Olson is now with the Marcus Autism Center at Emory University. This material is based on work supported by the National Institute on Deafness and Other Communication Disorders of the National Institutes of Health under awards R01DC007129 and R01DC012033, and by Autism Speaks under Innovative Technology for Autism Grant 2407. The content is solely the responsibility of the authors and does not necessarily represent the official views of the granting agencies or any other individual.

References

Acton, E.K. (2011). On gender differences in the distribution of um and uh. Penn Working Papers in Linguistics, 17(2), 1–9.

Adams, C., Green, J., Gilchrist, A., & Cox, A. (2002). Conversational behaviour of children with Asperger syndrome and conduct disorder. Journal of Child Psychology and Psychiatry, 43(5), 679–690.

Allwood, J., Nivre, J., & Ahlsén, E. (1990). Speech management: On the non-written life of speech. Nordic Journal of Linguistics, 13(1), 1–48.

American Psychiatric Association (2000). Diagnostic and statistical manual of mental disorders (4th ed. text revision, 943 p.). Washington, DC: American Psychiatric Publishing.

Arnold, J.E., et al. (2003). Disfluencies signal theee, um, new information. Journal of Psycholinguistic Research, 32(1), 25–36.

Bailey, K.G., & Ferreira, F. (2003). Disfluencies affect the parsing of garden-path sentences. Journal of Memory and Language, 49(2), 183–200.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practicaland powerful approach to significance testing. Journal of the Royal Statistical Society, Series B, 57(1): 289–300.

Berument, S.K., Rutter, M., Lord, C., Pickles, A., & Bailey, A. (1999). Autism screening questionnaire: Diagnostic validity. British Journal of Psychiatry, 175(5), 444–451.

Bishop, D.V.M. (2001). Pragmatic language impairment: A correlate of SLI, a distinct subgroup, or part of the autism continuum? In D.V.M. Bishop & L.B. Leonard, (Eds.), Speech and language impairments in children: Causes, characteristics, intervention and outcome (pp. 99–113). Philadelphia: Psychology Press.

Bishop, D.V.M. (2003). The children's communication checklist version 2 (CCC-2). San Antonio: Psychological Corporation.

Bishop, D.V.M., & Norbury, C.F. (2002). Exploring the borderlands of autistic disorder and specific language impairment: A study using standardised diagnostic instruments. Journal of Child Psychology and Pyschiatry, 43(3), 917–929.

Bishop, D.V.M., Whitehouse, A.J.O., Watt, H.J., & Line, E.A. (2008). Autism and diagnostic substitution: Evidence from a study of adults with a history of developmental language disorder. Developmental Medicine and Child Neurology, 50(5), 341–345.

Bock, J.K. (1982). Towards a cognitive psychology of syntax: Information processing contributions to sentence formulation. Psychological Review, 89(1), 1–47.

Bortfeld, H., Leon, S.D., Bloom, J.E., Schober, M.F., & Brennan, S.E. (2001). Disfluency rates in conversation: Effects of age, relationship, topic, role, and gender. Language and Speech, 44(2), 123–147.
Botting, N., & Conti-Ramsden, G. (2003). Autism, primary pragmatic difficulties, and specific language impairment: Can we distinguish them using psycholinguistic markers? Developmental Medicine and Child Neurology, 45(8), 515–524.
Brennan, S.E., & Clark, H.H. (1995). Conceptual pacts and lexical choice in conversation. Journal of Experimental Psychology: Learning, Memory, and Cognition, 22(6), 1482–1493.
Brennan, S.E., & Schober, M.F. (2001). How listeners compensate for disfluencies in spontaneous speech. Journal of Memory and Language, 44(2), 274–296.
Brown, R. (1973). A first language: The early stages (437 p.). Cambridge: Harvard University Press.
Capps, L., Kehres, J., & Sigman, M. (1998). Conversational abilities among children with autism and children with developmental delays. Autism, 2(4), 325–344.
Christenfeld, N. (1995). Options and ums. Journal of Language and Social Psychology, 13(2), 192–199.
Clark, H.H. (1994). Managing problems in speaking. Speech Communication, 15(3–4), 243–250.
Clark, H.H., & Fox Tree, J.E. (2002). Using uh and um in spontaneous speaking. Cognition, 84(1), 73–111.
Clark, H.H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. Cognition, 22(1), 1–39.
Cook, M., Smith, J., & Lalljee, M.G. (1974). Filled pauses and syntactic complexity. Language and Speech, 17(1), 11–16.
Cox, A., Klein, K., Charman, T., Baird, G., Baron-Cohen, S., Swettenham, J., . . ., Wheelwright, S. (1999). Autism spectrum disorders at 20 and 42 months of age: Stability of clinical and ADI-R diagnosis. Journal of Child Psychology and Psychiatry, 40(5), 719–732.
de Villiers, J. (2011). Use of um and uh in spontaneous speaking in autism spectrum disorder (pp. 101–110). In LACUS Forum XXXVI: Mechanisms of linguistic behaviour. http://www.lacus.org/volumes/36/205_devilliers_j.pdf
Ellis Weismer, S. (2013). Developmental language disorders: Challenges and implications of cross-group comparisons. Folia Phoniatrica et Logopaedia, 65(2), 68–77.
Engelhardt, P.E., Corley, M., Nigg, J.T., & Ferreira, F. (2010). The role of inhibition in the production of disfluencies. Memory and Cognition, 38(5), 617–628.
Engelhardt, P.E., Ferreira, F., & Nigg, J.T. (2013). Is the fluency of language outputs related to individual differences in intelligence and executive function? Acta Psychologica, 144(2), 424–432.
Fox Tree, J.E. (1995). The effects of false starts and repetitions on the processing of subsequent words in spontaneous speech. Journal of Memory and Language, 34(6), 709–738.
Fox Tree, J.E. (2001). Listeners' uses of um and uh in speech comprehension. Memory and Cognition, 29(2), 320–326.
Gioia, G.A., Isquith, P.K., Guy, S.C., & Kenworthy, L. (2000). Behavior rating inventory of executive function. Odessa, FL: Psychological Assessment Resources.
Gioia, G.A., Espy, K.A., & Isquith, P.K. (2003). Behavior rating inventory of executive function—Preschool version. Odessa, FL: Psychological Assessment Resources.
Gotham, K., Risi, S., Pickles, A., & Lord, C. (2007). The autism diagnostic observation schedule: Revised algorithms for improved diagnostic validity. Journal of Autism and Developmental Disorders, 37(4), 613–627.
Heeman, P.A., Lunsford, R., Selfridge, E., Black, L., & van Santen, J. (2010). Autism and interactional aspects of dialogue (pp. 249–252). Tokyo, Japan: In SIGdial.
Hus, V., Gotham, K., & Lord, C. (in press). Standardizing ADOS domain scores: Separating severity of social affect and restricted and repetitive behaviors. Journal of Autism and Developmental Disorders.
Irvine, C.A. (2014). Uh, um, and autism: Filler disfluencies in children with optimal outcomes from autism spectrum disorder. Master's thesis, University of Connecticut, Storrs, CT.
Jefferson, G. (1974). Error correction as an interactional resource. Language in Society, 3(2), 181–199.
Kenworthy, L., Yerys, B.E., Anthony, L.G., & Wallace, G.L. (2008). Understanding executive control in autism spectrum disorders in the lab and in the real world. Neuropsychology Review, 18(4): 320–338.
Kidd, C., White, K.S., & Aslin, R.N. (2011). Toddlers use speech disfluencies to predict speakers' referential intentions. Developmental Science, 14(4), 925–934.
Kim, S.H., Paul, R., Tager-Flusberg, H., & Lord, C. (2014). Language and communication in autism. In: F.R. Volkmar, S. Rogers, R. Paul, & K.A. Pelphrey Eds.), Handbook of autism and pervasive developmental disorders. Volume 1: Diagnosis, development, and brain mechanisms (4th ed., pp. 230–262). Hoboken, NJ: Wiley.
Kjelgaard, M., & Tager-Flusberg, H. (2001). An investigation of language impairment in autism: Implications for genetic subgroups. Language and Cognitive Processes, 16(2/3), 287–308.
Klin, A., Lang, J., Cicchetti, D.V., & Volkmar, F.R. (2000). Interrater reliability of clinical diagnosis and DSM-IV criteria for autistic disorder: Results of the DSM-IV autism field trial. Journal of Autism and Developmental Disorders, 30(2), 163–167.
Klin, A., McPartland, J.C., & Volkmar, F.R. (2005). Asperger syndrome. In: F.R. Volkmar, R. Paul, A. Klin, & D.J. Cohen (Eds.), Handbook of autism and pervasive developmental disorders (3rd ed., pp. 88–125). Hoboken: Wiley.
Klin, A., Saulnier, C.A., Sparrow, S.S., Cicchetti, D.V., Volkmar, F.R., & Lord, C. (2007). Social and communication abilities and disabilities in higher functioning individuals with autism spectrum disorders: The Vineland and the ADOS. Journal of Autism and Developmental Disorders, 37(4), 748–759.
Lai, C., Gorman, K., Yuan, J., & Liberman, M. (2007). Perception of disfluency: Language differences and listener bias (pp. 2345–2348). Antwerp, Belgium: In INTERSPEECH.
Lake, J.K., Humphreys, K.R., & Cardy, S. (2011). Listener vs. speaker-oriented aspects of speech: Studying the disfluencies of individuals with autism spectrum disorders. Psychonomic Bulletin and Review, 18(1), 135–140.
Lam, Y.G., & Yeung, S.S.S. (2012). Towards a convergent account of pragmatic language deficits in children with high-functioning autism: Depicting the phenotype using the Pragmatic Rating Scale. Research in Autism Spectrum Disorders, 6(2), 792–797.

Landa, R. (2000). Social language use in Asperger syndrome and high-functioning autism. In A. Klin, F.R. Volkmar, & S.S. Sparrow (Eds.), Asperger syndrome (pp. 125–155). New York: Guilford Press.

Landis, J.R., & Koch, G.G. (1977). The measurement of observer agreement for categorical data. Biometrics, 33(1), 159–174.

Lee, L.-C., DAvid, A.B., Rusyniak, J., Landa, R., & Newschaffer, C.J. (2007). Performance on the Social Communication Questionnaire in children receiving preschool special education services. Research in Autism Spectrum Disorders, 1(2), 126–138.

Levelt, W.J. (1983). Monitoring and self-repair in speech. Cognition, 14(1), 41–104.

Levelt, W.J. (1989). Speaking: From intention to articulation (584 p.). Cambridge: MIT Press.

Levelt, W.J., & Cutler, A. (1983). Prosodic marking in speech repair. Journal of Semantics, 2(2), 205–217.

Leyfer, O.T., Tager-Flusberg, H., Dowd, M., Tomblin, J.B., & Folstein, S.E. (2008). Overlap between autism and specific language impairment: Comparison of Autism Diagnostic Interview and Autism Diagnostic Observation Schedule scores. Autism Research, 1(5), 284–296.

Lickley, R.J., Shillcock, R.C., & Bard, E.G. (1991). Processing disfluent speech: How and when are disfluencies found? (pp. 1499–1502). Genova, Italy: In EUROSPEECH.

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 (2nd ed., pp. 195–225). New York: Wiley.

Lord, C., Risi, S., Lambrect, L., Cook, Jr., E.H., Leventhal, B.L., DiLavore, P.C., . . ., Rutter, M. (2000). The autism diagnostic observation schedule-generic: A standard measure of social and communication deficits associated with the spectrum of autism. Journal of Autism and Developmental Disorders, 30(3), 205–223.

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

Loucas, T., Charman, T., Pickles, A., Simonoff, E., Chandler, S., Meldrum, D., & Baird, G. (2008). Autistic symptomatology and language ability in autism spectrum disorder and specific language impairment. Journal of Child Psychology and Psychiatry, 49(11), 1184–1192.

Loveland, K.a., McEvoy, R.M., Tunali, B., & Kelley, M.L. (1990). Narrative story telling in autism and down's syndrome. British Journal of Developmental Psychology, 8(1), 9–23.

Lunsford, R. (2012). Towards improving dialogue coordination in spoken dialogue systems. Doctoral dissertation, Oregon Health & Science University, Portland, OR.

Lunsford, R., Heeman, P.A., Black, L., & van Santen, J.P. (2010). Autism and the use of fillers: Differences between 'um' and 'uh' (pp. 107–110). Tokyo, Japan: In DiSS-LPSS.

Maclay, H., & Osgood, C. (1959). Hesitation phenomena in spontaneous English speech. Word, 15(1), 19–44.

Martin, J.G., & Strange, W. (1968). The perception of hesitation in spontaneous speech. Perception and Psychophysics, 3(6), 427–438.

Miller, J.F., & Chapman, R.S. (1985). Systematic Analysis of Language Transcripts. Madison, WI: University of Wisconsin.

Oviatt, S. (1995). Predicting spoken disfluencies during human-computer interaction. Computer Speech and Language, 9(1), 19–35.

Paul, R., Augustyn, A., Klin, A., & Volkmar, F.R. (2005). Perception and production of prosody by speakers with autism spectrum disorders. Journal of Autism and Developmental Disorders, 35(2), 205–220.

Paul, R., Orlovski, S.M., Marcinko, H.C., & Volkmar, F.R. (2009). Conversational behaviors in youth with high-functioning ASD and Asperger syndrome. Journal of Autism and Developmental Disorders, 39(1), 115–125.

Peppé, S., McCann, J., Gibbon, F., O'Hare, A., & Rutherford, M. (2007). Receptive and expressive prosodic ability in children with high-functioning autism. Journal of Speech, Language, and Hearing Research, 50(4), 1015–1028.

Pinheiro, J., & Bates, D. (2000). Mixed-effects models in S and S-PLUS (530 p.). New York: Springer.

Ramberg, C., Ehlers, S., Nydén, A., Johansson, M., & Gillberg, C. (1996). Language and pragmatic functions in school-age children on the autistic spectrum. European Journal of Disorders of Communication, 31(4), 387–413.

Rayson, P., Leech, G., & Hodges, M. (1997). Social differentiation in the use of English vocabulary: Some analyses of the conversational component of the British National Corpus. International Journal of Corpus Linguistics, 2(1), 133–152.

Russell, R.L., & Grizzle, K.L. (2008). Assessing child and adolescent pragmatic language competencies: Towards evidence-based assessment. Clinical Child and Family Psychology Review, 11(1), 59–73.

Rutter, M., Bailey, A., & Lord, C. (2003). Social communication questionnaire (SCQ). Los Angeles: Western Psychological Services.

Sacks, H., Schegloff, E.A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. Language, 50(4), 696–735.

Schacter, S., Christenfeld, N., Ravina, B., & Bilous, F. (1991). Speech disfluency and the structure of knowledge. Journal of Personality and Social Psychology, 60(3), 362–367.

Semel, E., Wiig, E., & Secord, W. (2004). Clinical evaluation of language fundamentals preschool (2nd ed.). San Antonio: Psychological Corporation.

Semel, E., Wiig, E., & Secord, W. (2003). Clinical evaluation of language fundamentals (4th ed.). San Antonio: Psychological Corporation.

Shriberg, E.E. (1996). Disfluencies in Switchboard (pp. 11–14). Philadelphia, Pennsylvania: In ICSLP.

Shriberg, E.E., & Lickley, R.J. (1993). Intonation of clause-internal filled pauses. Phonetica, 50(3), 172–179.

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

Smith, V.L., & Clark, H.H. (1993). On the course of answering questions. Journal of Memory and Language, 32(1), 25–38.

Spitzer, R.L., & Siegel, B. (1990). The DSM-III field trial of pervasive developmental disorders. Journal of the American Academy of Child and Adolescent Psychiatry, 29(6), 855–862.

Stenström, A.-B. (1994). An introduction to spoken language interaction (256 p.). Harlow, England: Longman.
Stone, W.L., & Caro-Martinez, L.M. (1990). Naturalistic observations of spontaneous communication in autistic children. Journal of Autism and Developmental Disorders, 20(4), 437–453.
Su, Y.-S., Gelman, A., Hill, J., & Yajima, M. (2011). Multiple imputation with diagnostics (mi) in R: Opening windows into the black box. Journal of Statistical Software, 45(2), 1–31.
Suh, J., Eigsti, I.-M., Naigles, L., Barton, M., Kelley, E., & Fein, D. (2014). Narrative performance of optimal outcome children and adolescents with a history of an autism spectrum disorder (ASD). Journal of Autism and Developmental Disorders, 44(7), 1681–1694.
Swerts, M. (1998). Filled pauses as markers of discourse structure. Journal of Pragmatics, 30(4), 485–496.
Tager-Flusberg, H. (1996). Brief report: Current theory and research on language and communication in autism. Journal of Autism and Developmental Disorders, 26(2), 169–172.
Tager-Flusberg, H., & Joseph, R.M. (2003). Identifying neurocognitive phenotypes in autism. Philosophical Transactions of the Royal Society, Series B, 358: 303–314.
Tager-Flusberg, H., Paul, R., & Lord, C. (2005). Language and communication in autism. In F.R. Volkmar, R. Paul, A. Klin, & D.J. Cohen (Eds.). Handbook of autism and pervasive developmental disorders (3rd ed., pp. 335–364). Hoboken, NJ: Wiley.
Tager-Flusberg, H. Rogers, S., Cooper, J., Landa, R., Lord, C., Paul, R., ..., Yoder, P. (2009). Defining spoken language benchmarks and selecting measures of expressive language development for young children with autism spectrum disorders. Journal of Speech, Language, and Hearing Research, 52(3), 643–652.
Thordardottir, E.T., & Ellis Weismer, S. (2002). Content mazes and filled pauses in narrative language samples of children with specific language impairment. Brain and Cognition 43(2–3), 587–592.
Thurber, C., & Tager-Flusberg, H. (1993). Pauses in the narratives produced by autistic, mentally retarded, and normal children as an index of cognitive demand. Journal of Autism and Developmental Disorders, 23(2), 309–322.
Tomblin, B. (2011). Co-morbidity of autism and SLI: Kinds, kin and complexity. International Journal of Language and Communication Disorders, 46(2), 127–137.
Tottie, G. (2011). Uh and uhm as sociolinguistic markers in British English. International Journal of Corpus Linguistics, 16(2), 173–197.
van Santen, J.P.H., Prud'hommeaux, E.T., Black, L., & Mitchell, M. (2010). Computational prosodic markers for autism. Autism, 14(3), 215–236.
Volden, J., & Phillips, L. (2010). Measuring pragmatic language in speakers with autism spectrum disorders: Comparing the Children's Communication Checklist–2 and the Test of Pragmatic Language. American Journal of Speech-Language Pathology, 19(3), 204–212.
Volden, J., Coolican, J., Garon, N., White, J, & Bryson, S. (2009). Pragmatic language in autism spectrum disorder: Relationships to measures of ability and disability. Journal of Autism and Developmental Disorders, 39(2), 388–393.
Watanabe, M. (2001). The usage of fillers at discourse segment boundaries in Japanese lecture-style monologues (pp. 89–92). Edinburgh, Scotland: In DiSS.
Watanabe, M. (2002). Fillers as indicators of discourse segment boundaries in Japanese monologues (pp. 691–694). Aix-en-Provence, France: In Speech Prosody.
Watanabe, M., & Ishi, C.T. (2000). The distribution of fillers in lectures in the Japanese language. (pp. 167–170). Beijing, China: In ICSLP.
Wechsler, D. (2002). Wechsler preschool and primary scale of intelligence (3rd ed.). San Antonio: Psychological Corporation.
Wechsler, D. (2004). The Wechsler intelligence scale for children (4th ed.). San Antonio: Psychological Corporation.
Whitehouse, A.J.O., Barry, J.G., & Bishop, D.V.M. (2008). The broader language phenotype of autism: A comparison with specific language impairment. Journal of Child Psychology and Psychiatry, 48(8), 822–830.
Wilkins, D.P. (1992). Interjections as deictics. Journal of Pragmatics, 18(2–3), 119–158.