# Autism and Interactional Aspects of Dialogue

**Peter A. Heeman, Rebecca Lunsford, Ethan Selfridge, Lois Black,** and **Jan van Santen**
Center for Spoken Language Understanding
Oregon Health & Science University
heemanp@ohsu.edu

## Abstract

Little research has been done to explore differences in the interactional aspects of dialogue between children with Autistic Spectrum Disorder (ASD) and those with typical development (TD). Quantifying the differences could aid in diagnosing ASD, understanding its nature, and better understanding the mechanisms of dialogue processing. In this paper, we report on a study of dialogues with children with ASD and TD. We find that the two groups differ substantially in how long they pause before speaking, and their use of fillers, acknowledgments, and discourse markers.

## 1 Introduction

Autism Spectrum Disorders (ASD) form a group of severe neuropsychiatric conditions whose features can include impairments in reciprocal social interaction and in communication (APA, 2000). These impairments may take different forms, ranging from individuals with little or no communication to fully verbal individuals with fluent, grammatically correct speech. In this latter verbal group, shortcomings in communication have been noted, including using and processing social cues during conversations. This is no surprise, since negotiating a conversation requires many abilities, several of which are generally impaired in ASD, such as generating appropriate prosody (Kanner, 1943) and "theory of mind" (Baron-Cohen, 2000).

We make a distinction between transactional and interactional aspects of dialogue (Brown and Yule, 1983). The transactional aspect refers to message content and interactional focuses on expressing social relations and personal attitudes. In this paper, we focus on surface behaviors that speakers use to help manage the interaction, namely turn-taking, and the use of fillers, discourse markers, and acknowledgments. One advantage of these behaviors is that they do not require complete understanding of the dialogue, and thus lend themselves to automatic analysis. In addition, these behaviors are under the speaker's control and should be robust to what the other speaker is doing. We hypothesize that just as interactional aspects in general are affected in ASD, so are these surface behaviors. However, to our knowledge, little or no work has been done on this.

Investigating how the interactional aspects of dialogue are affected in ASD serves several purposes. First, it can help in the diagnostic process. Currently, diagnosing ASD is subjective. Objective measures based on dialogue interaction could improve the reliability of the diagnostic process. Second, it can help us refine the behavioral phenotypes of ASD, which is critical for progress on the basic science front. Third, it can help us refine therapy for people with ASD to address dialogue interaction deficits. Fourth, understanding what dialogue aspects are affected in high-functioning verbal children with ASD can help determine which aspects of dialogue are primarily social in nature. For example, do speakers use fillers to signal that there is a communication problem, or are fillers a symptom of it (cf. Clark and Fox Tree, 2002)?

In this paper, we report on a study of interactional aspects of dialogues between clinicians and children with ASD. The dialogues were recorded during administration of the Autism Diagnostic Observation Schedule (Lord et al., 2000), which is an instrument used to assist in diagnosing ASD. We compare the performance of these children with a group of children with typical development (TD).

## 2 Data

The data used in this paper was collected during administration of the ADOS on 22 TD children and 26 with ASD, ranging in age from 4 to 8 years old. The children with ASD were high-functioning and verbal. The speech of the clinician and child was transcribed into utterance-like units, with a start and an end time. Activities were annotated in a separate tier. The transcriptions included the punctuation marks '.', '!', and '?' to mark syntactically and semantically complete sen-

To appear in the 11th Annual SIGdial Meeting on Discourse and Dialogue, Tokyo, September 2010.

tences, and '>' to mark incomplete ones. As a single audio channel was used, the timing of overlapping speech was marked as best as possible. Each child on average said 2221 words, 574 utterances, and 316 turns.

## 3 Results

**Pauses between Turns:** We first examine how long children wait before starting their turn. We hypothesized that children with ASD would wait longer on average to respond, either because they are less aware of (a) the turn-taking cues, (b) the social obligation to minimize inter-turn pauses, or (c) they have a slower processing and response times. For this analysis, we look at all turns in which there is no overlap between the beginning of the child's turn and the clinician's speech. Data is available on 4412 pauses for the TD children and 5676 for the children with ASD. The grand means of the children's pauses are shown in Table 1 along with the standard deviations. The TD children's average pause length is 0.876s. For the children with ASD, it is 1.115s, 27.3% longer. This difference is significant, *a-priori* independent t-test $t=2.34$ (df=39), $p<.02$ one-tailed.

| | TD | | ASD | |
|---|---|---|---|---|
| all | 0.876 | (0.24) | 1.115 | (0.45) |
| after question | 0.748 | (0.25) | 1.005 | (0.40) |
| after non-question | 1.076 | (0.37) | 1.329 | (0.74) |

Table 1: Pauses before new turns.

We also examine the pauses following questions by the clinician versus non-questions. Questions are interesting as they impose a social obligation for the child to respond, and they have strong prosodic cues at their ending. We identified questions as utterances transcribed with a question mark, which might include rhetorical questions. After a non-question (e.g., a statement), the average pause is 1.076s for the TD children and 1.329s for children with ASD. This difference is not statistically significant by independent t-test, $t<1.6$, NS. After a question, the average pause is 0.748s for the TD children and 1.005s for the children with ASD, a significant difference by *a-priori* independent t-test $t=2.72$ (df=42), $p<.005$ one-tailed. The ASD children on average take 34.4% longer to respond. Thus, after a question, the difference between children with TD and ASD is more pronounced.

**Pauses by Activity:** The ADOS includes having the child engage in different activities. For this research, we collapse the activities into three types: *converse* is when there is no non-speech task; *describe* is when the child is doing a mental task, such as describing a picture; and *play* is when the child is interacting with the clinician in a play session. To better understand the difference between questions and non-questions, we examine the pauses in each activity (Table 2).

| | TD | | | | ASD | | | |
|---|---|---|---|---|---|---|---|---|
| | question | | non-ques. | | question | | non-ques. | |
| converse | 0.730 | 0.30 | 0.656 | 0.27 | 0.890 | 0.34 | 0.932 | 0.88 |
| describe | 0.853 | 0.44 | 0.879 | 0.37 | 1.056 | 0.51 | 1.282 | 1.21 |
| play | 0.720 | 0.34 | 1.825 | 0.78 | 1.289 | 1.51 | 1.887 | 1.37 |

Table 2: Pauses for each type of activity.

After a question, the TD children tend to respond with similar pauses in each activity (the differences in column 2 between activities are not significant by pairwise paired t-test, all t's$<1.6$, NS). After a question, the child has a social obligation to respond, and this does not seem to be overridden by whether there is a separate task they are involved in. Even after a non-question, conversants have a social obligation to keep the speaking floor occupied and so to minimize inter-utterance pauses (Sacks et al., 1974). However, as seen in the third column, the pauses are affected by the type of activity, and the differences are statistically significant by pairwise paired t-test, (df=21), two-tailed: converse-describe $t=2.24$, $p<.04$; describe-play $t=5.68$, $p<.0001$; converse-play $t=6.87$, $p<.0001$. The biggest difference is with *play*. Here, it seems that the conversants physical interaction lessens the social obligation of maintaining the speaking floor. These findings are interesting for social-linguistics as it suggests that the social obligations of turn-taking are altered by the presence of a non-speech task.

We next compare the children with ASD to the TD children. For the *converse* activity, we see that the children with ASD take longer to respond, after questions and non-questions. The difference after questions is significant by independent t-test, $t=1.74$ (df=46) $p<.05$, one-tailed, whereas the difference after non-questions is marginal, $t=1.47$ (df=28) $p<.08$. This result could be explained by the slower processing and response times associated with ASD.

Just as with the TD children, we see that after a non-question, the children with ASD take longer to respond when there is another task. The differences in pause lengths between *converse* and *play* are significant, by paired t-test, $t=2.89$ (df=23) p<.009, two-tailed. The difference between *describe* and *play* is marginal, t=2.03 (df=25) p<.06, and there was no significant difference between *converse* and *describe*, t<1, NS.

After a question, the children with ASD take longer to respond when there is another task, especially for *play*, although the pairwise differences in pause length between activities are not significant. This suggests that the children with ASD become distracted when there is another task, and so become less sensitive to either the question prosody or the social obligation of questions.

**Fillers:** We next examine the rate of fillers, at the beginning of turns, beginning of utterances, and in the middle of utterances. We look at these contexts individually as fillers can serve different roles, such as turn-taking, stalling for time or as part of a disfluency, and their role is correlated to their position in a turn. The rates are reported in Table 3, along with the total number of fillers within each category. Interestingly, the rate of 'uh' between children with TD and ASD is similar for all positions (independent t-test, all g's<1, NS).

| | uh | | | | um | | | |
|---|---|---|---|---|---|---|---|---|
| | TD | | ASD | | TD | | ASD | |
| turn init. | 1.70% | 112 | 1.84% | 159 | 3.86% | 243 | 1.65% | 146 |
| utt. init. | 1.31% | 43 | 1.20% | 33 | 2.29% | 73 | 0.52% | 10 |
| utt. medial | 0.25% | 103 | 0.31% | 137 | 1.03% | 492 | 0.21% | 123 |

Table 3: Rate of fillers.

The more interesting finding, though, is in the usage of 'um'. Children with ASD use it significantly less than the TD children in every position, from $1/2$ the rate in turn-initial position to $1/5$ in utterance-medial position, independent two-tailed t-test: turn initial t=2.74 (df=38), p<.01; utterance initial t=2.53 (df=31), p<.02; and utterance medial t=3.94 (df=24), p<.001.

| | TD | | ASD | |
|---|---|---|---|---|
| converse | 1.76% | 569 | 0.56% | 190 |
| describe | 1.15% | 115 | 0.33% | 31 |
| play | 0.96% | 124 | 0.45% | 58 |

Table 4: Use of 'um' by activity.

We also examined the overall usage of 'um' in each activity (Table 4). The TD children use 'um' more often in each activity than the children with ASD, and the differences are statistically significant by independent two-tailed t-test: converse t=3.62 (df=29), p<.002; describe t=2.83 (df=27), p<.01; play t=2.42 (df=33), p<.03. This result supports the robustness of the findings about 'um'.

Many researchers have speculated on the role of 'um' and 'uh'. In recent work, Clark and Fox Tree (2002) argued that they signal a delay, and that 'um' signals more delay than 'uh'. They view both as linguistic devices that are planned for, just as any other word is. Our work suggests that 'um' and 'uh' arise from different cognitive processes, and that the process that accounts for 'uh' is not affected by ASD, while the process for 'um' is.[1]

**Acknowledgments:** We next look at the rate of acknowledgments: single word utterances that are used to show agreement or understanding. Thus, the use of acknowledgments requires awareness of the other person's desire to ensure mutual understanding. As the corpus did not have these words explicitly marked, we identify a word as an acknowledgment if it meets the following criteria: (a) it is one of the words listed in Table 5 (based on Heeman and Allen, 1999); (b) it is first in the speaker's turn; and (c) it does not follow a question by the clinician. The TD children used acknowledgments in 17.42% of their turns that did not follow a question, while the children with ASD did this only 13.39% of the time (Table 5), a statistically significant difference by *a-priori* independent t-test t=1.78 (df=46), p<.05 one-tailed.

| | TD | | ASD | |
|---|---|---|---|---|
| total | 17.42% | 568 | 13.39% | 459 |
| yeah | 7.49% | 248 | 5.87% | 215 |
| no | 2.78% | 78 | 2.06% | 63 |
| mm-hmm | 2.06% | 75 | 1.07% | 35 |
| mm | 0.99% | 29 | 1.35% | 42 |
| ok | 1.87% | 65 | 0.83% | 27 |
| yes | 0.92% | 32 | 0.88% | 32 |
| right | 0.14% | 5 | 0.23% | 8 |
| hm | 0.73% | 21 | 0.69% | 20 |
| uh-huh | 0.44% | 15 | 0.42% | 17 |

Table 5: Use of acknowledgments.

**Discourse Markers:** We next examine discourse markers, which are words such as 'well' and 'oh' that express how the current utterance relates to the discourse context (Schiffrin, 1987). We classified a word as a discourse marker if it was the first word in an utterance and is one of the words in Table 6 (Heeman and Allen, 1999). As shown in Table 6, the children with ASD use discourse markers significantly less than the TD children in both conditions by *a-priori* independent, one-tailed t-test: turn-initial t=3.24 (df=43) p<.002; utterance-initial t=4.01 (df=44) p<.0001.

[1] In Lunsford et al. (2010) we investigate the rate and length of pauses after 'uh' and 'um'. In addition, we verified the t-tests using Wilcoxon rank sum tests.

As can be seen, most of the difference is in the use of ‘and’. The data for the other discourse markers was sparse, so we compared ‘and’ against all of the others combined. The decreased usage of ‘and’ in the ASD children is statistically significant for both conditions by *a-priori* independent, one-tailed t-test: turn-initial t=4.47 (df=30), p<.0001; utterance-initial t=3.79 (df=43), p<.0002. There is little difference in the use of all of the other discourse markers combined, and the difference is not statistically significant.

| | Turn Initial | | | | Utterance Initial | | | |
|---|---|---|---|---|---|---|---|---|
| | TD | | ASD | | TD | | ASD | |
| all | 19.2% | 1290 | 12.8% | 1196 | 28.7% | 2053 | 19.4% | 1330 |
| and | 10.7% | 731 | 5.0% | 471 | 19.5% | 1419 | 12.0% | 844 |
| then | 0.6% | 38 | 1.0% | 89 | 1.5% | 97 | 1.4% | 79 |
| but | 2.1% | 144 | 1.3% | 113 | 3.6% | 238 | 2.7% | 194 |
| well | 2.2% | 143 | 2.7% | 271 | 1.1% | 74 | 1.2% | 79 |
| oh | 2.0% | 135 | 1.8% | 160 | 1.0% | 67 | 1.3% | 68 |
| so | 1.2% | 75 | 0.7% | 60 | 1.6% | 129 | 0.7% | 49 |
| wait | 0.2% | 9 | 0.2% | 21 | 0.2% | 17 | 0.2% | 15 |
| actually | 0.2% | 15 | 0.1% | 11 | 0.2% | 12 | 0.0% | 2 |
| not and | 8.5% | 559 | 7.8% | 725 | 9.2% | 634 | 7.4% | 486 |

Table 6: Use of discourse markers.

The use of ‘and’ is also lower in each activity for the ASD children (Table 7), a significant difference by *a-priori* independent one-tailed t-test: converse t=3.00 (df=41), p<.003; describe t=4.79 (df=38), p<.0001, play t=4.07 (df=30), p<.0002.

| | TD | | ASD | |
|---|---|---|---|---|
| converse | 13.36% | 1139 | 7.95% | 755 |
| describe | 21.77% | 587 | 10.76% | 339 |
| play | 12.97% | 424 | 5.18% | 221 |

Table 7: Use of ‘and’ in each activity.

One explanation for the decreased usage of ‘and’ and not the other discourse markers might be that, of all the discourse markers, ‘and’ seems to have the least meaning. It simply signifies that there is some continuation between the new speech and the previous context. This might make it difficult for children with ASD to learn its use. A second explanation is that the children with ASD are using ‘and’ correctly, but simply do not produce as many utterances that are related to the previous context (cf. Bishop et al., 2000).

## 4 Conclusion

In this paper, we examined a number of interactional aspects of dialogue in the speech of children with ASD and TD. We found that children with ASD have a lower rate of the filler ‘um’, acknowledgments, and the discourse marker ‘and.’ We also found that in certain situations, they take longer to respond. These deficits might prove useful for improved diagnosis of ASD. We also found that children with ASD have a lower rate of ‘um’ but not of ‘uh’, and that only the discourse marker ‘and’ seems to be affected. This might prove useful for both better understanding the nature of ASD as well as better understanding the role of these phenomena in dialogue. Although the results reported in this work are preliminary, they do show the potential of our approach. More work is needed to ensure that our automatic identification of turn-taking events, discourse markers, and acknowledgments is correct and to explore alternate explanations for the results that we observed.


## Acknowledgments

Funding gratefully received from the National Institute of Heath under grants IR21DC010239 and 5R01DC007129, and the National Science Foundation under IIS-0713698. The views herein are those of the authors and reflect the views neither of the funding agencies.


## References


American Psychiatric Association, Washington DC, 2000. *Diagnostic and Statistical Manual of Mental Disorders, 4th Edition, Text Revision (DSM-IV-TR).*

S. Baron-Cohen. 2000. Theory of mind and autism: A review. In L. M. Glidden, editor, *International Review of Research in Mental Retardation*, volume 23: Autism, pages 170–184. Academic Press.

D. Bishop et al. 2000. Conversational responsiveness in specific language impairment: Evidence of disproportionate pragmatic difficulties in a subset of children. *Development and Psychopathology*, 12(2):177–199.

G. Brown and G. Yule. 1983. *Discourse Analysis*. Cambridge University Press.

H. Clark and J. Fox Tree. 2002. Using uh and um in spontaneous speaking. *Cognition*, 8:73–111.

P. Heeman and J. Allen. 1999. Speech repairs, intonational phrases and discourse markers: Modeling speakers’ utterances in spoken dialog. *Computational Linguistics*, 25(4):527–572.

L. Kanner. 1943. Autistic disturbances of affective content. *Nervous Child*, 2:217–250.

C. Lord et al. 2000. The autism diagnostic observation schedule-generic: a standard measure of social and communication deficits associted with the spectrum of autium. *Journal of Austism Developmental Disorders*, 30(3):205–223, June.

R. Lunsford et al. 2010. Autism and the use of fillers: differences between ‘um’ and ‘uh’. In *5th Workshop on Disfluency in Spontaneous Speech*, Tokyo.

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

D. Schiffrin. 1987. *Discourse Markers*. Cambridge University Press, New York.