Proceedings of DiSS-LPSS Joint Workshop 2010,
25-26 September 2010, Tokyo, Japan, pp. 107-110.

Autism and the use of fillers: Differences between 'um' and 'uh'

Rebecca Lunsford, Peter A. Heeman, Lois Black, and Jan van Santen

Center for Spoken Language Understanding, OHSU, Beaverton, OR USA

lunsforr@ohsu.edu

Abstract

Little research has been done to explore differences in the use of the fillers 'um' and 'uh' between children with Autistic Spectrum Disorder (ASD) and those with typical development (TD). Quantifying any differences could aid in diagnosing ASD, understanding its nature, and better understanding the mechanisms involved in dialogue processing. In this paper, we report on a study of dialogues between clinicians and children with ASD or TD, finding that the two groups of children differ substantially in their use of 'um' but not 'uh'. This suggests that these two fillers result from different cognitive processes.

Index Terms: disfluencies, fillers, autism

1. Introduction

Autism Spectrum Disorders (ASD) form a group of severe neuropsychiatric conditions which are characterized by impaired reciprocal social interaction and communication, repetitive behaviors and restricted interests. In terms of communication, 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, nevertheless shortcomings in communication have been noted. Among these are shortcomings in 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 contextually appropriate responses [1] and "theory of mind" [2].

In this paper we examine the use of 'um' versus 'uh', as these have been posited to have a social function during dialogue. Research that quantifies differences in the use of fillers[1] can serve several purposes. First, it can help in the diagnostic process. Currently, the instruments used to aid diagnosis of ASD are fundamentally subjective making it difficult to rate reliably. Objective measures based on the interactional aspects of dialogue could improve the reliability of the diagnostic process. Second, understanding what dialogue aspects are affected by ASD could help refine the behavioral phenotypes of ASD, which is critical for progress on the basic science front. Third, understanding what aspects of dialogue are affected by ASD in high-functioning verbal children with ASD (for whom their social communication abilities are affected, but not their basic verbal abilities) can help determine which aspects of dialogue interaction are primarily social in nature. It is this third purpose which is of particular interest to this workshop.

Previous work on spontaneous speech by Clark and Fox Tree presented the "Filler-as-word" hypothesis [3]. This hypothesis states that fillers are words used to announce the initiation of what is expected to be a delay in speaking, with 'uh' signaling a minor delay and 'um' signaling a major delay. Clark and Fox Tree suggest that, when using these fillers, speakers fulfill a specific social obligation; that of informing the listener of upcoming delays in the speaker's speech. If fillers do fulfill the social use posited in this work, one might anticipate that children with ASD would have difficulty in using fillers as expected, but that typically developing (TD) children would use fillers as described by Clark and Fox Tree. Specifically, it seems likely that children with ASD would use less fillers than TD children, and that the 'um's would not necessarily indicate more or longer pauses than 'uh's.

In previous work we examined dialogues between clinicians and children with ASD, as well as a group of TD children [4]. We compared interactional aspects of their dialogue including turn-taking behavior, filler use, discourse markers, and acknowledgments, finding substantial differences between the two groups in many categories. In particular, we found that children with ASD differed from TD children only in how often they use 'um'; however, surprisingly they did not differ in how they often they used 'uh'. We suggested that 'um' and 'uh' result from different cognitive processes, with the process responsible for 'um' affected by ASD, but not the process responsible for 'uh'.

Our previous work leads us to now hypothesize that 'um' might result from a social process (intentional signal to the hearer), while 'uh' might result from a self-directed process (un-intentional signal to self). In this paper, we augment our previous analyses by also exploring whether 'um' and 'uh' are used to differentially signal upcoming speech delays. We expect that TD children will have a higher ratio of pauses and longer pauses after 'um', as compared to 'uh', but we do not expect to see significant differences between 'um' and 'uh' for the children with ASD. Based on our previous findings showing a difference in the rate of 'um's, but not 'uh's, for TD children versus children with ASD, we hypothesize that the two groups will have similar rates and lengths of pauses after 'uh', but that they will differ for 'um'. In this paper, we compare the two groups ratios of fillers to total words, and also their ratios and incidence at turn-initial, utterance-initial, and utterance-medial positions. We also compare the ratios of fillers while the children engage in three different activities; conversation, description, and play. Finally, we analyze the ratio and length of pauses following fillers.

2. Related Work

2.1. Role of Fillers

Many researchers have speculated on the role of 'um' and 'uh'. Swerts suggests that fillers can carry information about topical units in a dialogue, with stronger breaks in the discourse more likely to co occur with fillers than weaker breaks [5]. In addition, Swerts found that fillers at strong breaks are more likely to be preceded and followed by pauses. Goldman-Eisler suggested

[1] In this paper we use the term "filler" to refer to 'uh' and 'um'. In other works they are often referred to as "filled pauses".

that fillers signal a speaker's word-searching problems [6], and Stenström found they can function as turn-holders [7].

In recent work, Clark and Fox Tree [3] hypothesized that 'um' and 'uh' are words used by a speaker to signal a delay in the speech, with 'um' signaling a longer delay than 'uh'. They viewed both fillers as linguistic devices that are planned for, just as any other word is, and that they meet the roles described in the previous paragraph by implicature. For example, fillers allow the speaker to imply their intent to hold the floor by signaling an expected delay in speaking, and to imply uncertainty by signaling a delay in formulating an utterance. To support this hypothesis, Clark and Fox Tree compared speakers use of 'um' and 'uh', in a study primarily using the London-Lund corpus of face-to-face conversations [3]. They found that (a) speakers use fillers most often near utterance boundaries, primarily 'um's, (b) 'um's are more likely to be followed by a pause than 'uh's, and (c) pauses following 'um's are longer than pauses following 'uh's.

More recent work by O'Connell and Kowal [8] questioned this hypothesis. In this work, they analyzed the speech of professional news and radio interviewers, along with Hillary Clinton, while engaged in interviews regarding Hillary's upcoming book. They found that, for these practiced public speakers, the ratio of 'uh's to 'um's was much higher then that found in Clark and Fox Tree's work, and that fillers did not reliably signal delays in speaking, especially for the interviewers (as compared to the interviewee). All of this previous work treats 'um' and 'uh' as similar phenomena.

2.2. ASD and Private Speech

Children with ASD are known to have deficits in both executive functioning and the use of social speech. In contrast, recent work has suggested that children with ASD have relatively unimpaired private (i.e., self-directed) speech [9]. High-functioning children with ASD were found to use private speech at the same rate as children with typical development. In addition, in contrast to TD children, children with ASD were more likely to get items correct when talking versus when silent [9]. This works suggest that children with ASD use private speech to bolster their executive functioning and improve task performance. Thus, if a filler were to arise from a cognitive process responsible for producing private speech, we could expect children with ASD to use fillers in a manner similar to TD children. In contrast, if fillers arise from a cognitive process responsible for the use of social speech, we would expect children with ASD to use fillers less often and less effectively than TD children.

3. Data

The data used in this paper was collected during administration of the Autism Diagnostic Observation Schedule [10] in which 48 children, 22 with TD and 26 with ASD, engaged in different activities with a clinician. The activities consisted of having a conversation, describing a wordless picture or book, and playing with toys. The children ranged in from 4 to 8. The clinician and child's speech was annotated using praat [11]. Speech was segmented into communication-units [12], and transcribed with a start and an end time. C-units were transcribed with '.', '!', and '?' to mark syntactically and semantically complete sentences, and '>' to mark incomplete ones. Pauses, both between units and intra-unit, were annotated in a manner similar to the London-Lund corpus, along with their start and end times. Finally, the activity was annotated in a separate tier.

4. Results

There were 21,836 utterances consisting of 106,592 words available for analysis. All children produced both fillers, with the exception of one TD child who produced only 'um's (no 'uh's). In total, the children produced 1,674 fillers, 1087 'um's and 587 'uh's.

4.1. Ratio of Fillers

First, we examined the ratio of fillers as a percentage of total words for each child. As illustrated in Figure 1,[2] TD children's average ratio of fillers was 2.03%, significantly higher than 1.15% for children with ASD, Wilcoxon rank-sum test, z=2.16, p<.02. one-tailed. Looking separately at 'um' and 'uh', it is clear that this difference is accounted for primarily by the ASD group's significantly lower ratio of 'um's, 0.48% vs. 1.50% for the TD group, Wilcoxon rank-sum test, z=3.69, p<.001, one-tailed. The two group's ratio of 'uh's, 0.53% and 0.66%, were not significantly different by Wilcoxon rank-sum test, z<1.4, NS. Additionally, comparing within group, TD children's ratio of 'um's was significantly higher than their ratio of 'uh's, Wilcoxon signed-rank test, N=21, z=3.34, p<.001, one-tailed, but no significant difference was found between the ratios of 'um's and 'uh's for the children with ASD, N=25, z<1.6, NS.

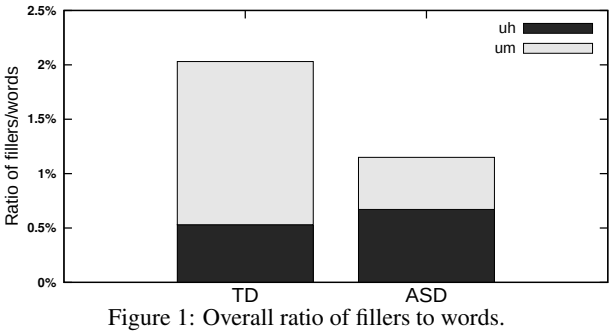

Figure 1: Overall ratio of fillers to words.

We next examined the ratio of fillers, factored by whether the filler appears at the beginning of a new turn (turn-initial), the beginning of an utterance within a turn (utterance-initial), or elsewhere in an utterance (utterance-medial). The ratios are reported in Figure 2 and Table 1.[2] Comparing between groups, we see that the average ratio of 'uh' for both TD children and those with ASD is similar for all positions, with no significant differences by Wilcoxon rank-sum test, all z's<1.0, NS. For 'um', the children with ASD had a significantly lower ratio than that of the TD children in every position by Wilcoxon rank-sum test, one-tailed: turn-initial z=2.89, p<0.002; utterance-initial z=3.53, p<0.001; and utterance-medial z=4.01, p<0.001.

Comparing within group, we see that TD children's ratio of 'um' was higher than their ratio of 'uh' for all positions, significantly so for turn-initial (N=21, z=2.57, p<.006) and utterance-medial (N=20, z=3.88, p<.001), and marginally so for utterance-initial (N=18, z=1.61, p<.06), by Wilcoxon signed-rank test, one-tailed. In contrast, the children with ASD had higher ratios of 'uh' than 'um' in all positions, significantly so for utterance-initial (N=16, z=2.43, p<.01), marginally so for utterance-medial positions (N= 21, z=1.67, p<.06), but showed no significant difference in the turn-initial position, N=21, z<1, NS, by Wilcoxon signed-rank test, one-tailed.

[2]Results regarding the incidence of fillers are replicated from [4]. For this work we replaced parametric analyses with non-parametric and added within-group analyses where appropriate.

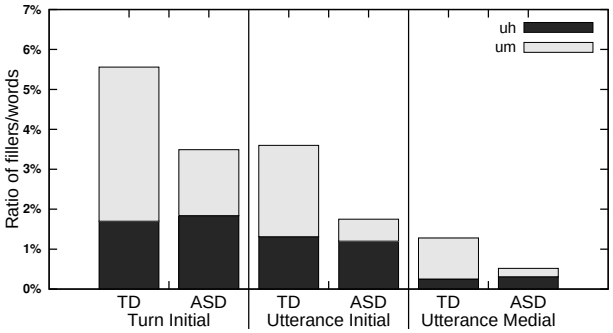

Figure 2: Ratio of filler within each position.

	um				uh			
	TD		ASD		TD		ASD	
	ratio	n	ratio	n	ratio	n	ratio	n
turn-initial	3.86%	21	1.65%	22	1.70%	18	1.84%	23
utterance-initial	2.29%	17	0.52%	6	1.31%	12	1.20%	15
utterance-medial	1.03%	21	0.21%	21	0.25%	20	0.31%	24

Table 1: Ratio of fillers within each position and the number of children with fillers in that category.

Table 1 details the data shown in Figure 2, and also includes the number of children with fillers in each category. Both TD children and children with ASD average a higher ratio of 'um's at utterance boundaries (i.e., turn-initial and utterance-initial) than utterance-medial. This agrees with Clark and Fox Tree on the use of 'um'. However, looking at the number of children who produced fillers, it is clear that the two groups differ in terms of utterance initial 'um's, with only 6 of 26 children with ASD producing an 'um' at the beginning of an utterance compared to 17 of 22 TD children, a significant difference by Fisher exact test, $p<.001$, one-tailed.

We also examined the overall usage of fillers in each activity. The results are reported in Figure 3.[2] The TD children use 'um' far more often in each activity than the children with ASD, with all differences statistically significant by Wilcoxon rank-sum test, one-tailed: conversation $z=3.62$, $p<0.001$; description $z=2.95$, $p<0.002$; play $z=2.24$, $p<0.02$. No significant differences were found for 'uh', all z's<1.3, NS.

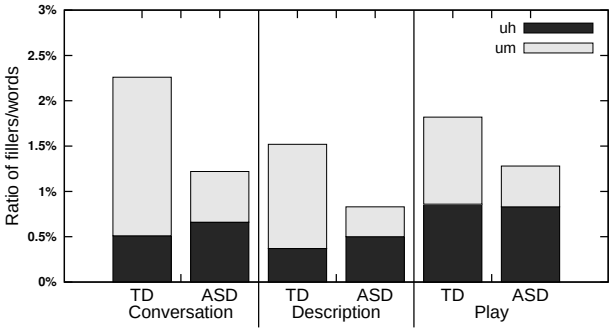

Figure 3: Ratio of fillers within each activity.

From the above results, it is clear that, although children with ASD have similar rates of usage for 'uh', their rates of 'um' are much lower than TD children, and that this difference is robust regardless of the position of the filler within the utterance or the child's activity.

4.2. Pauses after Fillers

Next we examined the ratio of fillers that were followed by pauses. For this measure, data were excluded for fillers which were followed by the clinician's speech, as the length of the pause was not under the child's control. For these analyses a total of 1466 fillers were available, of which 660 were followed by non-zero length pauses. Table 2 shows the ratio of fillers followed by pauses for each group, as well as the number of children in that group who produced any pauses after that filler. The TD group paused after 49.22% of their 'um's, significantly more often than 33.01% for the ASD group, Wilcoxon rank-sum test, $z=1.79$, $p<.04$, one-tailed. The two group's ratio of pauses after 'uh's were not significantly different by Wilcoxon rank-sum test, $z<1$, NS. In addition, the TD children had a significantly higher ratio of pauses after 'um's than after 'uh's by Wilcoxon signed-rank test, $N=21$, $z=2.36$, $p<.01$, one-tailed, but the children with ASD did not, $N=23$, $z<1.0$, NS. In fact, only 16 out of 24 (67%) children with ASD ever paused after 'um', whereas all but one of the TD children did, a significant difference by Fisher exact test, $p<.02$, one-tailed.

	TD		ASD	
	ratio	(n/N)	ratio	(n/N)
um	49.22%	(21/22)	33.01%	(16/24)
uh	32.92%	(17/21)	38.39%	(23/26)

Table 2: Ratio of fillers followed by pauses and the proportion of children (n/N) who produced pauses.

We next compare the length of pauses following fillers, including 0-length pauses. As shown in Table 3, columns 1 and 2, TD children averaged 0.61 seconds pause after 'um's, marginally longer than the ASD group's average of 0.46 seconds, Wilcoxon rank-sum test, $z=1.31$, $p<.10$, one-tailed. There was no significant difference between the two group's average pause length after 'uh's, $z<0.6$, NS. In addition, TD children had significantly longer pauses after 'um' versus 'uh', by Wilcoxon signed-rank test, $N=21$, $z=2.36$, $p<.01$, one-tailed, but the children with ASD did not, $N=24$, $z<1.3$, NS. This supports our hypothesis that TD children will average longer pauses after 'um' than after 'uh', but children with ASD will not.

	including 0-length		excluding 0-length	
	TD	ASD	TD	ASD
um	0.61	0.46	1.15	1.41
uh	0.36	0.35	1.00	1.15

Table 3: Mean length of pauses (in seconds) following fillers, both including and excluding 0-length pauses.

In order to allow comparison with O'Connell and Kowal[8], we also looked at only those fillers that were followed by a non-zero length pause, as shown in Table 3, columns 3 and 4. For this measure, the two groups showed no significant differences in the length of pauses after 'um' (TD:1.15s vs ASD:1.41s), or after 'uh' (TD:1.00s vs ASD:1.15s), by Wilcoxon rank-sum test, all z's<1.4. However, comparing within group, both groups did have significantly longer pauses after 'um' than after 'uh' by Wilcoxon signed-rank test, one-tailed: TD $N=16$, $z=1.91$, $p<.03$; ASD $N=13$, $z=2.59$, $p<.01$. This suggests that, for those children who produced pauses after both fillers, 'um' signals a longer delay than 'uh'. However, only 1/2 the children with ASD, and 2/3 of the TD children met this criteria.

5. Conclusion

In this paper we compared the use of fillers by children with ASD versus TD children. Our results for TD children replicated Clark and Fox Tree's findings for adults. That is, TD children paused significantly more often, and longer, after 'um'' as compared to 'uh'. In addition, TD children used more 'um's than 'uh's at utterance boundaries. It is particularly interesting that this pattern of filler use is already established in young TD children (i.e., aged 4-8).

Comparing the two groups of children, we found that TD children and children with ASD differed significantly in their use of 'um', but not in the use of 'uh'. Children with ASD used 'um' less often, were less likely to pause after 'um' and had shorter pauses after 'um'. Essentially, children with ASD displayed none of the regular, directional differences between ''um' and 'uh' seen for the TD children.

One might posit that these differences arise because children with ASD lack the ability to learn the differing social functions of 'uh' and 'um' (i.e., that of signaling a minor versus major delay), thus use them interchangeably. However, this explanation does not account for the low incidence of children with ASD who use 'um's at the beginning of a utterance, or the higher ratio of 'uh's in utterance initial and utterance medial positions. That is, if using the two fillers interchangeably, the children would use both 'uh' and 'um' equally, rather than preferring 'uh' in some instances.

We offer two explanations for the differences in 'um' usage between the two groups of children. The first is that, because children with ASD have an impairment in executive functioning, it is difficult for them to accurately anticipate and appropriately signal upcoming delays. The second is that the two result from different cognitive processes: 'uh' from an internally focused process similar to that for private speech, in which the speaker uses 'uh' to assist himself; and 'um' from an externally focused process, in which the speaker uses 'um' to assist the listener. The second explanation better accounts for the greater use of 'uh' by children with ASD, because it places 'uh' within the realm of self-directed private speech, which has been shown to be especially effective for children with ASD. However, given that children with ASD have known deficits both in executive functioning and in processing and using social cues, more work will need to be done to tease out the source of these differences. Interestingly, our second explanation gives an alternative account for O'Connell and Kowal's results: it is possible that experienced public speakers, who typically strive to minimize disfluencies, are able to suppress their outwardly focused 'um's, but are less successful at suppressing inwardly focused 'uh's.

Although, at this point, we can only speculate as to the source of the differential use of 'um's by children with ASD, these finding are nevertheless immediately relevant to several purposes. First, the unique pattern of filler use by children with ASD can be used to assist in diagnosis, specifically by looking for a tendency to use 'uh' more than 'um' at utterance initial and medial positions. Second, this work provides additional insight into dialogue processing in ASD, leading toward a finer understanding of which dialogue skills are affected by ASD. Third, by illustrating the similarities and differences in use of fillers between TD children and those with ASD, we gain a greater understanding of the mechanisms involved in dialogue processing, with particular insights into what aspects of dialogue may be more social in nature.

6. Acknowledgements

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 do not reflect the views of the funding agencies.

7. References

[1] C. Adams, J. Green, A. Gilchrist, and A. Cox, "Conversational behaviour of children with asperger syndrome and conduct disorder," *Journal of Child Psychology and Psychiatry*, vol. 43, no. 5, pp. 679–690, 2002.

[2] S. Baron-Cohen, "Theory of mind and autism: A review," in *International Review of Research in Mental Retardation*, L. M. Glidden, Ed. New York: Academic Press, 2000, vol. 23: Autism, pp. 170–184.

[3] H. H. Clark and J. E. Fox Tree, "Using uh and um in spontaneous speaking." *Cognition*, vol. 84, no. 1, pp. 73–111, May 2002.

[4] P. A. Heeman, R. Lunsford, E. Selfridge, L. Black, and J. van Santen, "Autism and interactional aspects of dialogue," in *SIGdial*, 2010.

[5] M. Swerts, "Filled pauses as markers of discourse structure," *Journal of Pragmatics*, vol. 30, no. 4, pp. 485–496, October 1998.

[6] F. Goldman-Eisler, *Psycholinguistics: Experiments in spontaneous speech.* Academic Press, 1968.

[7] A.-B. Stenström, *An introduction to spoken language interaction.* Longman, 1994.

[8] D. O'Connell and S. Kowal, "Uh and um revisited: Are they interjections for signaling delay?" *Journal of Psycholinguistic Research*, vol. 34, no. 6, pp. 555–576, November 2005.

[9] A. Winsler, B. Abar, M. A. Feder, C. D. Schunn, and D. A. Rubio, "Private speech and executive functioning among high-functioning children with autistic spectrum disorders," *Jounal of Autism & Developmental Disorders*, vol. 37, pp. 1617–1635, 2007.

[10] C. Lord, S. Risi, L. Lambrecht, E. Cook, B. Leventhal, P. DiLavore, A. Pickles, and M. Rutter, "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*, vol. 30, no. 3, pp. 205–223, June 2000.

[11] P. Boersma, "Praat, a system for doing phonetics by computer." *Glot International*, vol. 5, no. 9/10, pp. 341–345, 2001.

[12] "Segmenting utterances into c-units,", Salt Software, LLC [Online]. Available: http://www.saltsoftware.com/training/ transciption/ TransConvSummary.pdf, Accessed 2010.