Computational Analysis of Backchannel Usage and Overlap Length in Autistic Children

Grace O. Lawley[1], **Peter A. Heeman**[1], **Steven Bedrick**[2]
[1]*Computer Science and Engineering*
[2]*Department of Medical Informatics and Clinical Epidemiology*
Oregon Health & Science University, Portland, OR, USA

Abstract

Backchanneling (e.g., *right*, *okay*, *uhhuh*) during a dialogue signals that a person is engaged and following along with what is being said. Although backchannels often overlap with other utterances, they are not interpreted as an attempt to take the floor when used successfully. Limited work has been done on investigating the frequency and overlap length of backchannels in the language of Autistic children. After controlling for age, sex, and IQ, we found that Autistic children used significantly less backchannels than their Typically Developing (TD) peers. Additionally, we found that Autistic children were less likely than TD children to produce a backchannel with a greater overlap length.

1 Introduction

Providing feedback to your conversational partner in the form of backchanneling is a pervasive component of verbal communication. A backchannel is a short utterance (e.g., *mmhmm*, *yes*, *uhhuh*) said by person A while person B continues to have the floor (Levinson and Torreira, 2015). Although these brief utterances do not contribute new meaning to the dialogue, they still contribute important pragmatic information; by using a backchannel, a person is signaling that they are engaged and following along but that they also understand the other person is not ready to yield the floor. Backchannels sometimes (but not always) overlap other utterances (Levinson and Torreira, 2015).

Deficits in backchanneling ability could lead to miscommunications or problems related to turn-taking. An extended pause before a backchannel could cause the backchannel to be interpreted as negative rather than positive (e.g., an excessive pause before saying *okay*). Starting a backchannel too close to the end of the other speaker's utterance could be interpreted as an attempt to take the floor (Schegloff, 2000).

Autism Spectrum Disorder (ASD) is a developmental condition characterized by difficulties with social communication (American Psychiatric Association, 2013). These difficulties often include problems with pragmatic language, with Autistic[1] children frequently having difficulties with conversational reciprocity, such as turn taking and filler usage (Baltaxe and D'Angiola, 1992; Paul et al., 2009). Computational methods have been used before to successfully capture differences between Autistic and Typically Developing (TD) children for fillers and mazes (Parish-Morris et al., 2017; Salem et al., 2021; Lawley et al., 2022).

Another way these difficulties could manifest is in backchannel usage. While some work has been done on examining differences in backchannel usage in ASD before (Heeman et al., 2010; Lunsford et al., 2012), previous investigations have been limited by small sample size, few female participants, and lack of controlling for participant-level variables such as age, sex, and intellectual ability. Furthermore, to our knowledge, previous work has not examined backchanneling in combination with overlap length.

In this paper we investigate whether Autistic children use backchannels at different rates their TD peers using a multivariate approach that allows us to control for potential confounding participant-level variables such as age, sex, and IQ. Since Autistic children frequently have difficulties with pragmatic language skills, we hypothesize that the ASD group will use less backchannels than the TD group. We also investigate whether group differences in backchannel rates are affected by whether the backchannel is an overlapping utterance and the length of the overlap (if any). Assuming that producing an overlapping-backchannel requires better

[1]We are using identity-first language (i.e., Autistic children) here instead of person-first language (i.e., children with Autism) as the former is the current preference among many Autistic individuals (Brown, n.d.).

Proceedings of the First ICARD Workshop, pages 17–23
September 12, 2023

turn-taking abilities than producing a backchannel that does not overlap, we hypothesize that the ASD group will produce less overlapping-backchannels and the ones they do produce will have a shorter overlap length.

2 Methods

2.1 Dataset

Analyses were performed on language samples of 116 ASD children and 65 TD children between 4 to 15 years old. All were native English speakers and had an IQ ≥ 70. Intellectual ability was estimated using the Wechsler Intelligence Scale for Children, fourth edition (WISC-IV; Wechsler, 2003) for children 7 years and older; for children younger than 7 years old, the Wechsler Preschool and Primary Scale of Intelligence, third edition (WPPSI-III; Wechsler, 2002) was used instead. Overall language ability and pragmatic and structural language skills were measured using the Children's Communication Checklist, version 2 (CCC-2; Bishop, 2003). A complete summary of the demographic and clinical sample characteristics for all 181 children is reported in Appendix A.

2.2 Language Samples

Analyses were performed on transcribed Autism Diagnostic Observation Schedule (ADOS) sessions (Lord et al., 2000). All children were administered the ADOS-2, Module 3. Scoring was completed using the revised algorithms (Gotham et al., 2009). Sessions were transcribed by a team of trained transcribers according to modified Systematic Analysis of Language Transcript (SALT) guidelines (Miller and Iglesias, 2012).

Transcribers were instructed to include one of the following five punctuation marks at the end of each utterance: '.', '?', '!', '>', '∧'. Transcribers used '>' for abandoned utterances and '∧' for interrupted utterances. Spans of overlapping speech were surrounded by angled brackets: $< >$. Four ADOS activities were selected for this analysis: *Emotions*; *Social Difficulties and Annoyance*; *Friends, Relationships and Marriage*; *Loneliness*. These activities were chosen because of their conversational structure and similarities to naturalistic dialogue.

2.3 Backchannels

For each child, we calculated the total number of utterances that were backchannels. We considered an utterance to be a backchannel if it (1) appeared in the following, predefined list: *mmhmm*, *yes*, *ok*, *uhhuh*, *right*, *yeah*, *yep*; (2) was not the first utterance of the transcript; (3) did not follow a question (i.e., its predecessor utterance, as defined in section 2.4, was not a question). Creation of the predefined list of backchannels was informed by spelling conventions followed by our transcription team for words that commonly occur in natural conversation. These spelling conventions were strictly followed during transcription. We omitted utterances that immediately followed a question to avoid catching instances where words such as *yes* and *yeah* were used as an affirmative reply to a question. Overall, there were a total of 1,187 backchannels: 753 *yeah*; 223 *mmhmm*; 75 *yes*; 49 *yep*; 43 *ok*; 34 *uhhuh*; 10 *right*.

There were a total of 1,807 utterances that satisfied criteria (1) and (2) but not (3). To test the validity of our rule of omitting these utterances, we took a random sample of 200 utterances from this subset (100 for each diagnostic group) and manually checked each. Of the 200 random utterances, 2 were false positives (i.e., were backchannels) and 198 were true negatives (i.e., were not backchannels), giving us a false positive rate of 0.01.

2.4 Overlap Length

For a given utterance, we defined the overlap length as the amount (in seconds) that it overlaps with its predecessor utterance. We followed the process detailed by Lunsford et al. (2016) to identify the predecessors of each utterance. Given an utterance u, let u' be the previous utterance said by the same speaker. Let w be the most recent utterance said by the second speaker (i.e., start time of $w <$ start time of u). Whichever of u' and w has the later end time is the predecessor of u. Not every utterance is a predecessor utterance and a single utterance can be the predecessor for multiple utterances. The initial utterance in a transcript will not have a predecessor. An example of predecessor utterances in practice is shown in Figure 1.

After identifying the predecessors for each utterance, we can proceed with calculating the overlap length (if any) of each utterance and its predecessor. We first subtracted the end time of the predecessor from the start time of the utterance. If the resulting value is positive (i.e., a pause or gap), the overlap length is 0. If instead the resulting value is negative (i.e., an overlap), the overlap length is the absolute

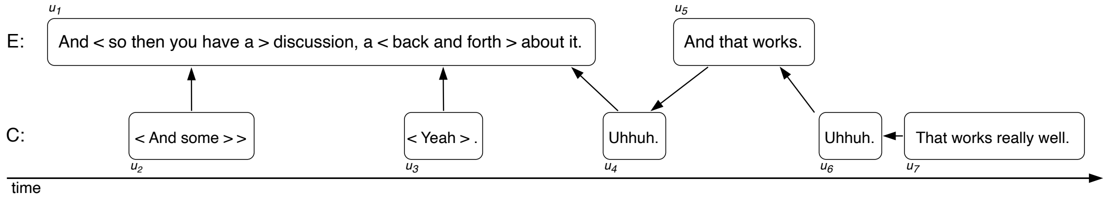

Figure 1: Example of predecessor utterances. Arrows point towards the predecessor of a given utterance. Abbreviations: E = Examiner; C = Child.

value of this number.

2.5 Overlapping-backchannels

We defined an overlapping-backchannel as an utterance that is (1) a backchannel and (2) overlaps its predecessor utterance by more than 200 ms. By this definition, overlapping-backchannels are a subset of backchannels. We used a cutoff of 200 ms to account for any overlaps that can be attributed to reaction time delays (Fry, 1975; Levinson and Torreira, 2015).

2.6 Statistical Analysis

We first compared the rates of both backchannels (total backchannels / total utterances) and overlapping-backchannels (total overlapping-backchannels / total utterances) between the ASD and TD groups without incorporating additional participant-level variables. Since normality assumptions were not met (Shapiro-Wilk Normality test; $p < .001$), we used the nonparametric Wilcoxon-Mann-Whitney test to compare groups. The dependent variable was backchannel or overlapping-backchannel rate and the independent variable was diagnosis (ASD; TD). We used rank-biserial correlations (r_{rb}; interpretation: small = 0.10-0.29, medium = 0.30-0.49, large = 0.50-1.0) to calculate effect size (Cureton, 1956; Wendt, 1972).

Next, to investigate group differences in backchannel rates while also taking into account the participants' age, sex, and IQ as well as overlap length, we fit a mixed effects logistic regression model. The binary response variable was created as follows: with the data formatted as one utterance per row, each backchannel was coded as 1 and every other utterance was coded as 0. A per-participant random intercept was included since each participant was associated with multiple utterances. The primary predictor variable was diagnosis (ASD; TD). The other predictor variables included were participants' age, sex, and IQ and the utterance overlap length. Additionally, an interaction term between diagnosis and overlap length was included. All continuous variables were transformed into z-scores prior to model estimation.

Lastly, we repeated our second experiment but for just the overlapping-backchannels. To create the binary response variable, each overlapping-backchannel was coded as 1 and every other utterance was coded as 0. We did not include a diagnosis and overlap length interaction term in this model since the results of Analysis of Variance (ANOVA) showed that the inclusion of an interaction term did not significantly contribute to the model.

All analyses were performed using the statistical programming language R (R Core Team, 2020). The lme4 package (Bates et al., 2015) was used to create the mixed effects logistic regression models.

3 Results

Figure 2 shows the distribution of both the backchannel and overlapping-backchannel rates within each diagnostic group.

The median and interquartile range (IQR) values and the results from the Wilcoxon-Mann-Whitney tests for backchannel and overlapping-backchannel rates are reported in Table 1. There was a significant difference in backchannel usage between the ASD and TD groups ($p = .001$; small effect size: $r_{rb} = .298$). The ASD group used less backchannels than the TD group overall (ASD = .025 [.011, .042] $<$ TD = .039 [.022, .066]). For overlapping-backchannels, there was also a significant group difference ($p < .001$; medium effect size: $r_{rb} = .397$), with the ASD group producing less overlapping-backchannels than the TD group (ASD = .001 [.000, .007] $<$ TD = .009 [.000, .018]).

Next, the results of the mixed effects logistic regression model for backchannel usage are reported in Table 2. A significant group difference in backchannel usage was still found after adjusting

	ASD	TD	U	p	r_{rb}
backchannels	.025 [.011, .042]	.039 [.022, .066]	2645.5	.001	.298
overlapping-backchannels	.001 [.000, .007]	.009 [.000, .018]	2273.0	$<$ .001	.397

Table 1: Backchannel and overlapping-backchannel usage rates by diagnostic group.

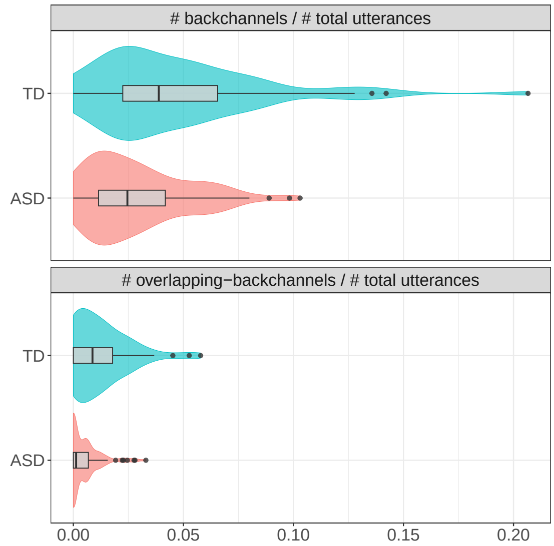

Figure 2: Distributions of backchannel and overlapping-backchannel rates by diagnosis. The x-axis (shared by both plots) is the proportion of backchannels or overlapping-backchannels said by a child. Behind the boxplots are violin plots. Violin plots are mirrored kernel density plots, where wider areas correspond to a higher density of observations.

for age, sex, IQ, and overlap length ($\chi^2 = -3.212$, $P = .001$, Table 2). As before, the ASD had a lower backchannel rate than the TD group. There was no significant effect on backchannel rate of participant age, sex, or IQ. Overlap length significantly contributed to backchannel rate ($\chi^2 = 9.651$, $P < .001$), with overlap length increasing the likelihood that an utterance is a backchannel. There was also a significant interaction between diagnosis and overlap length ($\chi^2 = -2.216$, $P = .027$), with the ASD group being less likely to produce a backchannel as the overlap length increases.

Lastly, the results of the mixed effects logistic regression model for overlapping-backchannels are reported in Table 3. For this model, inclusion of an interaction term between diagnosis and overlap length did not significantly contribute to the model so the interaction was left out. After controlling for age, sex, IQ, and overlap length, a significant group difference in overlapping-backchannel usage remained ($\chi^2 = -3.990$, $P < .001$), with the ASD group again using less backchannels than the TD group. The age, sex, and IQ of the participants had no significant effect on overlapping-backchannel rate. The overlap length significantly effected the likelihood that an utterance was an overlapping-backchannel (χ^2 = 19.496, $P < .001$), irrespective of participant's age, sex, IQ, or diagnosis. In other words, the longer the overlap, the more likely that an utterance was an overlapping-backchannel.

	Log-odds	S.E.	χ^2	$P(\chi^2)$
(Intercept)	-3.233	0.140		
Dx			-3.212	0.001
ASD	-0.496	0.154		
Sex			0.008	0.994
Male	0.001	0.149		
Age	0.025	0.067	0.375	0.708
IQ	-0.027	0.072	-0.379	0.705
Overlap	0.215	0.022	9.651	$<$ 0.001
Dx:Overlap			-2.216	0.027
ASD:Overlap	-0.082	0.037		

Table 2: Mixed effects logistic regression model predicting likelihood of a backchannel utterance.

	Log-odds	S.E.	χ^2	$P(\chi^2)$
(Intercept)	-4.845	0.210		
Dx			-3.990	$<$ 0.001
ASD	-0.949	0.238		
Sex			-0.765	0.444
Male	-0.176	0.229		
Age	-0.075	0.109	-0.690	0.490
IQ	-0.038	0.117	-0.328	0.743
Overlap	0.424	0.022	19.496	$<$ 0.001

Table 3: Mixed effects logistic regression model predicting likelihood of an overlapping-backchannel utterance.

4 Conclusion

We investigated backchannel usage differences in language samples of Autistic and TD children and examined whether differences were associated with participant age, sex, IQ, or utterance length. We found that not only did the ASD group use backchannels and overlapping-backchannels at a significantly lower rate than the TD group, but that these differences were robust to participants' age, sex, and intellectual abilities.

We also explored the effect of overlap length between an utterance and its predecessor utterance. After accounting for diagnosis, age, sex, and IQ, utterances were more likely to be backchannels the more they overlapped with their corresponding predecessor utterance. The diagnostic group and overlap length interaction significantly effected the likelihood an utterance would be a backchannel, with the ASD group being less likely than the TD group to produce a backchannel with a greater overlap length.

These results suggest that Autistic children use backchannels less than TD children and that this difference is affected by whether the backchannel overlaps and how long the overlap is. This could indicate that the TD group is more skilled at timing backchannels since they produced more overlapping utterances than the ASD group. Our next steps including further refining our method of calculating overlap length and investigating the potential underlying language processes associated with this difference.

Ethics Statement

This study was approved by the Oregon Health & Science University IRB (Protocol #531) and all research was performed in accordance with their relevant guidelines and regulations.

Acknowledgements

This work was supported in part by the National Institute on Deafness and Other Communication Disorders of the NIH under Awards R01DC012033 (PI: Dr. E. Fombonne) and R01DC015999 (PIs: Dr. S. Bedrick & G. Fergadiotis).

References

American Psychiatric Association. 2013. *Diagnostic and Statistical Manual of Mental Disorders*, 5th edition.

Christiane A. M. Baltaxe and Nora D'Angiola. 1992. Cohesion in the discourse interaction of autistic, specifically language-impaired, and normal children. *Journal of Autism and Developmental Disorders*, 22(1):1–21.

Douglas Bates, Martin Mächler, Ben Bolker, and Steve Walker. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1):1–48.

Dorothy V. M. Bishop. 2003. *The Children's Communication Checklist, version 2 (CCC-2)*. Pearson, London.

Lydia Brown. n.d. Identity-First Language. Autistic Self Advocacy Network (ASAN). https://autisticadvocacy.org/about-asan/identity-first-language/ Last accessed on 2023-05-10.

Edward E. Cureton. 1956. Rank-biserial correlation. *Psychometrika*, 21(3):287–290.

D.B. Fry. 1975. Simple Reaction-Times to Speech and Non-Speech Stimuli. *Cortex*, 11(4):355–360.

Katherine Gotham, Andrew Pickles, and Catherine Lord. 2009. Standardizing ADOS Scores for a Measure of Severity in Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*, 39(5):693–705.

Peter A. Heeman, Rebecca Lunsford, Ethan Selfridge, Lois Black, and Jan P. H. van Santen. 2010. Autism and Interactional Aspects of Dialogue. In *Proceedings of the SIGDIAL 2010 Conference*, pages 249–252, Tokyo, Japan. Association for Computational Linguistics.

Grace O. Lawley, Steven Bedrick, Heather MacFarlane, Jill K. Dolata, Alexandra C. Salem, and Eric Fombonne. 2022. "Um" and "Uh" Usage Patterns in Children with Autism: Associations with Measures of Structural and Pragmatic Language Ability. *Journal of Autism and Developmental Disorders*.

Stephen C. Levinson and Francisco Torreira. 2015. Timing in turn-taking and its implications for processing models of language. *Front. Psychol.*, 6(731).

Catherine Lord, Susan Risi, Linda Lambrecht, Edwin H. Cook, Bennett L. Leventhal, Pamela C. DiLavore, Andrew Pickles, and Michael Rutter. 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.

Rebecca Lunsford, Peter A. Heeman, and Emma Rennie. 2016. Measuring Turn-Taking Offsets in Human-Human Dialogues. In *Proc. Interspeech 2016*, pages 2895–2899.

Rebecca Lunsford, Peter A. Heeman, and Jan P. H. van Santen. 2012. Interactions Between Turn-taking Gaps, Disfluencies and Social Obligation. In *Proc. Interspeech 2012*, pages 607–610.

J. Miller and A. Iglesias. 2012. SALT: Systematic analysis of language transcripts [Research version]. *Middleton, WI: SALT Software*.

Julia Parish-Morris, Mark Y. Liberman, Christopher Cieri, John D. Herrington, Benjamin E. Yerys, Leila Bateman, Joseph Donaher, Emily Ferguson, Juhi

Pandey, and Robert T. Schultz. 2017. Linguistic camouflage in girls with autism spectrum disorder. *Molecular Autism*, 8(1):48–60.

Rhea Paul, Stephanie Miles Orlovski, Hillary Chuba Marcinko, and Fred Volkmar. 2009. Conversational Behaviors in Youth with High-functioning ASD and Asperger Syndrome. *Journal of Autism and Developmental Disorders*, 39(1):115–125.

R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Alexandra C. Salem, Heather MacFarlane, Joel R. Adams, Grace O. Lawley, Jill K. Dolata, Steven Bedrick, and Eric Fombonne. 2021. Evaluating atypical language in autism using automated language measures. *Scientific Reports*, 11(1):10968.

Emanuel A. Schegloff. 2000. Overlapping talk and the organization of turn-taking for conversation. *Language in Society*, 29(1):1–63.

David Wechsler. 2002. WPPSI-III: Wechsler Preschool and Primary Scale of Intelligence - 3rd ed. *San Antonio, TX: Psychological Corporation*.

David Wechsler. 2003. WISC-IV: Wechsler Intelligence Scale for Children. *San Antonio, TX: Psychological Corporation*.

Hans W. Wendt. 1972. Dealing with a common problem in social science: A simplified rank-biserial coefficient of correlation based on the U statistic. *European Journal of Social Psychology*, 2(4):463–465.

A Appendix

See Table 4 ("Demographic and clinical sample characteristics") on page 7.

	ASD *(n = 116, 97 males)*				TD *(n = 65, 37 males)*				
	min	*max*	*mean*	*s.d.*	*min*	*max*	*mean*	*s.d.*	*p*
Age in years	4.54	15.6	10.06	2.81	4.21	14.5	8.22	2.83	$<.001$
IQ	72	138	102.17	15.81	90	147	116.94	12.37	$<.001$
ADOS SA	3	19	9.17	3.49	0	8	0.95	1.47	$<.001$
ADOS RRB	0	8	3.59	1.54	0	2	0.45	0.64	$<.001$
ADOS Total	7	24	12.76	3.74	0	10	1.40	1.79	$<.001$
CCC-2 Pragmatic	1.5	10.8	4.96	1.70	7.5	15.8	12.05	1.73	$<.001$
CCC-2 Structural	1	12	7.01	2.30	8.5	15	11.73	1.57	$<.001$
CCC-2 GCC	45	103	75.12	11.04	87	143	115.18	12.09	$<.001$

Table 4: Demographic and clinical sample characteristics. Abbreviations: ADOS = Autism Diagnostic Observation Schedule; SA = Social Affect; RRB = Restricted and Repetitive Behavior; CCC-2 = Children's Communication Checklist, version 2; GCC = Global Communication Composite.