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, Oregon, USA

ICARD Workshop @ SIGdial & INLG 2023
September 12th, 2023

OHSU

What are backchannels?

What are backchannels?

- Short utterances, e.g. *mmhmm, yes*, *uhhuh*
  - Said during a conversation by person A while person B continues to have the floor
  - Sometimes but not always overlaps other utterances
- Contribute important pragmatic information
  - Person A is engaged and following along but also understands that person B is not ready to yield the floor

What are backchannels?

- Short utterances, e.g. *mmhmm, yes*, *uhhuh*
  - Said during a conversation by person A while person B continues to have the floor
  - Sometimes but not always overlaps other utterances
- Contribute important pragmatic information
  - Person A is engaged and following along but also understands that person B is not ready to yield the floor
- Deficits in backchanneling ability could lead to miscommunications or problems related to turn-taking
  - An extended pause before a backchannel —> could be interpreted as a negative response (e.g. an excessive pause before *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

In this paper

In this paper

(1) Investigate whether Autistic children use backchannels at different rates than their TD peers using a multivariate approach and control for potential confounding participant level variables (age, sex, and IQ)

- Hypothesis: ASD group will use less backchannels than the TD group

In this paper

(1) Investigate whether Autistic children use backchannels at different rates than their TD peers using a multivariate approach and control for potential confounding participant level variables (age, sex, and IQ)

- Hypothesis: ASD group will use less backchannels than the TD group

(2) Investigate whether group difference in backchannel rates are affected by whether a backchannel is an overlapping utterance and the length of the overlap (if any)

- Hypothesis: Assuming that producing an overlapping-backchannel requires better turn-taking abilities than producing a backchannel that does not overlap, ASD group will produce less overlapping-backchannels and the ones they do produce will have a shorter overlap length.

Dataset

- 116 ASD children, 65 TD children
- 4 to 15 years old
- IQ $\geq$ 70

Dataset

- 116 ASD children, 65 TD children
- 4 to 15 years old
- IQ $\geq$ 70

Language samples

- Transcribed ADOS-2, Module 3 sessions
- Four activities included in this analysis
  - (1) *Emotions*; (2) *Social Difficulties and Annoyance*; (3) *Friends, Relationships and Marriage*; (4) *Loneliness*
- Transcribers included end-of-sentence punctuation — **. ? !**
  - Abandoned utterances marked with **>**
  - Interrupted utterances marked with **^**
  - Spans of overlapping text surrounded with **< >**

ADOS = Autism Diagnostic Observation Schedule

Backchannels

- For each child, we calculated the total number of utterances that were backchannels

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 was not a question)

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 was not a question)
- Overall, there were a total of 1,187 backchannels

753 yeah	43 ok
223 mmhmm	34 uhhuh
75 yes	10 right
49 yep	

Overlap Length

- For a given utterance, we defined the **overlap length** as the amount (in seconds) that it overlaps with its **predecessor** utterance

Overlap Length

- For a given utterance, we defined the **overlap length** as the amount (in seconds) that it overlaps with its **predecessor** utterance

- Following Lunsford et al. (2016), we identified predecessors of each utterance as follows:

  > 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

Overlap Length

- For a given utterance, we defined the **overlap length** as the amount (in seconds) that it overlaps with its **predecessor** utterance

- Following Lunsford et al. (2016), we identified predecessors of each utterance as follows:

  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.

Overlap Length

- For a given utterance, we defined the **overlap length** as the amount (in seconds) that it overlaps with its **predecessor** utterance

- Following Lunsford et al. (2016), we identified predecessors of each utterance as follows:

  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.

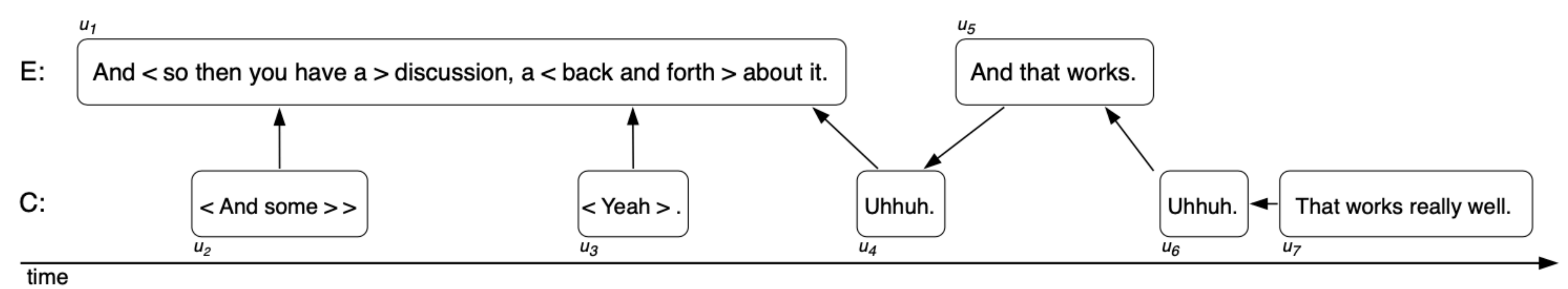

Figure 1: Example of predecessor utterances. Arrows point towards the predecessor of a given utterance. Abbreviations: E = Examiner; C = Child.

Overlapping-backchannel

- For each child, we calculated the total number of utterances that were overlapping-backchannels

Overlapping-backchannel

- For each child, we calculated the total number of utterances that were overlapping-backchannels

- We defined an **overlapping-backchannel** as an utterance that is
  - (1) a backchannel
  - (2) overlaps its predecessor utterance by more than 200 ms.

- By this definition, overlapping-backchannels are a subset of backchannels.

Overlapping-backchannel

- For each child, we calculated the total number of utterances that were overlapping-backchannels
- We defined an **overlapping-backchannel** as an utterance that is
  - (1) a backchannel
  - (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).

Experiment 1

(1) Compare backchannel and overlapping-backchannel rates between ASD and TD groups without incorporating participant level variables

- Backchannel rate = # backchannels / total utterances
- Overlapping-backchannel rate = # overlapping-backchannels / total utterances
- Wilcoxon-Mann-Whitney tests

Experiment 1 — Results

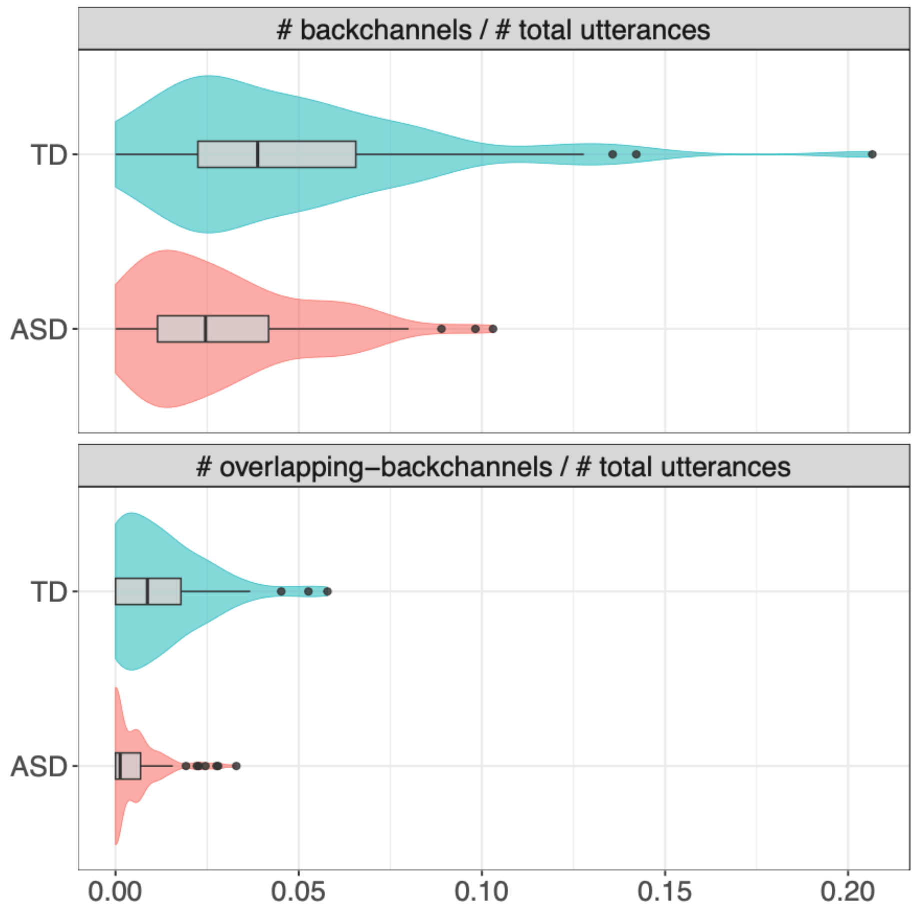

Figure 2: Distributions of backchannel and overlapping-backchannel rates by diagnosis. The x-axis (shared by both plots) is the proportion of backchannels or overlapping-backchannels said by a child. Behind the boxplots are violin plots. Violin plots are mirrored kernel density plots, where wider areas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

	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.

- 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),
  - ASD group produced less overlapping-backchannels than the TD group (ASD = .001 [.000, .007] < TD = .009 [.000, .018]).

Experiment 2

(2) Compare backchannel usage rates while taking into account age, sex, IQ, and overlap length

- Input data formatted as one utterance per row
- Mixed effects logistic regression model with binary response variable (1 if utterance is a backchannel, 0 if not)
- A per-participant random intercept was included since each participant was associated with multiple utterances.
- Primary predictor variable = diagnosis (ASD; TD).
- Other predictor variables = participants' age, sex, and IQ and the utterance overlap length.
- Included an interaction term between diagnosis and overlap length was included
- All continuous variables were transformed into z-scores prior to model estimation

Experiment 2 — Results

	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.

- A significant group difference in backchannel usage was still found after adjusting for age, sex, IQ, and overlap length (χ^2= -3.212, P = .001).
    - 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 (χ^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 (χ^2 = -2.216, P = .027), with the ASD group being less likely to produce a backchannel as the overlap length increases.

Experiment 3

(3) Repeat second experiment but this time for overlapping-backchannels

- Mixed effects logistic regression model with binary response variable (1 if utterance is a overlapping-backchannel, 0 if not)
- 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.

Experiment 3 — Results

	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.

- After controlling for age, sex, IQ, and overlap length, a significant group difference in overlapping-backchannel usage remained (χ^2= −3.990, P < .001)
  - 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.

Conclusion

- After controlling for age, sex, and IQ, ASD group used backchannels and overlapping-backchannels at a significantly lower rate than the TD group
- 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 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
  - Could indicate that the TD group is more skilled at timing backchannels since they produced more overlapping utterances than the ASD group
- Future work may include further refining our method of calculating overlap length and investigating the potential underlying language processes associated with this difference

Thank you

Computational Analysis of Backchannel Usage and Overlap Length in Autistic Children

Grace O. Lawley, Peter A. Heeman, Steven Bedrick

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).

Grace Olive Lawley
PhD Candidate, Computer Science & Engineering
Oregon Health & Science University
Portland, Oregon, USA
https://grace.rbind.io

OHSU