Lexical entrainment on target words during task-oriented interaction in children with and without autism spectrum disorder

Joanna Kruyt
Institute of Informatics,
Slovak Academy of Sciences/
in Bratislava, Slovakia
FIIT, Slovak Technical University
in Bratislava, Slovakia
joanna.kruyt@savba.sk

Katarína Polónyiová and **Daniela Ostatníková**
Academic Research Center for Autism,
Institute of Physiology, Faculty of Medicine,
Comenius University in Bratislava, Slovakia

Štefan Beňuš
Institute of Informatics, Slovak Academy of Sciences in Bratislava, Slovakia/
Constantine the Philosopher University in Nitra, Slovakia

Abstract

One widely observed strategy that interlocutors use to facilitate mutual understanding during dialogue is the repetition of each other's words, or lexical entrainment. Despite being well-researched, the underlying mechanisms of the phenomenon are debated. Specifically, the role of social factors and theory of mind are contested. This study aimed to investigate the role of theory of mind and neurotype on lexical entrainment. We recruited children with and without autism spectrum disorder, asked them to complete a collaborative task with an adult, and measured how often they entrained to the experimenter on "dispreferred" terms. We administered tests to measure IQ, executive functioning, and theory of mind for each child. Our results suggest that neither neurotype (i.e. autistic or neurotypical) nor theory of mind score predict entrainment, but that increased executive functioning difficulty predicts lower entrainment. Additionally, gender seems to influence entrainment. Theoretical implications of these results are discussed.

1 Introduction

During dialogue, two interlocutors need to collaborate to ensure that they understand one another. Mutual understanding can be achieved through several strategies. One of these strategies is the tendency of interlocutors to behave more similarly over time. This tendency is often referred to as entrainment, though other terms such as alignment, convergence, or synchrony are also used. This paper focuses on entrainment at the lexical level, i.e. on similarity in word choice.

Though entrainment has been widely observed, the exact psychological mechanisms underlying the phenomenon are debated. Specifically the role of higher-order cognition, in particular mentalising or "theory of mind" (ToM), is a topic of discussion. Some theories of entrainment postulate that it is an automatic process that occurs through priming mechanisms (interactive alignment hypothesis, Pickering and Garrod (2004, 2013)), while another theory is based on the idea of audience design: interlocutors tailor their utterances to whomever they are talking to, and take into account their "common ground" or mutually shared knowledge, which required perspective-taking and ToM skills (common ground/audience design account, Clark and Marshall (1978); Clark and Murphy (1982)). Yet another theory hypothesises that entrainment occurs because interlocutors aim to emphasise or minimise social differences between themselves and the person they are interacting with (Communication Accommodation Theory, Giles et al. (1991)).

In other words, the role of social and higher-order cognitive factors in entrainment is unclear. One way to elucidate the role of these external and internal factors, is by investigating entrainment in a group of people that exhibits both social and cognitive differences compared to the general population. Autism spectrum disorder (ASD) is often said to involve both of these: individuals with ASD report struggling with friendships and romantic relationships more than their neurotypical (NT) peers (e.g. Bossaert et al., 2015; Taheri et al., 2016), and ASD is associated with differences in ToM processing (e.g. Baron-Cohen, 2000; Baron-Cohen et al., 1985; Tager-Flusberg, 2007). Investigating entrainment in individuals with ASD can further inform us about the relationship between ToM and entrainment. This study aims to compare lexical entrainment in children with ASD and their NT peers, to characterise any potential between-group

Proceedings of the 27th Workshop on the Semantics and Pragmatics of Dialogue, August, 16–17, 2023, Maribor.

differences and to examine the role of ToM in entrainment.

2 Previous work

Entrainment in individuals with ASD has been investigated on several levels including syntax and lexical choice. Research suggests that individuals with ASD show similar levels of syntactic entrainment to individuals without ASD, both in experimental settings (Allen et al., 2011; Slocombe et al., 2013) and more naturalistic conversations (Hopkins et al., 2016). In terms of lexical entrainment, with which the present study is concerned, results from existing studies appear somewhat less consistent.

Lexical entrainment in individuals with ASD has been investigated using different methodologies: some studies focus on entrainment on target words, while others focus on overall lexical entrainment. Entrainment on target words is typically measured in a collaborative card-placing task during which an experimenter uses uncommon or "dispreferred" words to describe objects. Whether individual with ASD also adopt this dispreferred term is taken as a measure of entrainment to the experimenter. When such paradigms are used to measure lexical entrainment in individuals with ASD, results typically suggest that individuals with and without ASD do not show different entrainment patterns (e.g. Slocombe et al., 2013; Branigan et al., 2016; Hopkins et al., 2016). Importantly, conversations during such tasks are usually highly constrained, with predictable turn-taking and short turns. Du Bois et al. (2014) even refer to the speech during such structured tasks as "serial monologue" (p. 436) rather than dialogue, highlighting how such structured tasks do not resemble naturalistic, interactive conversation.

Rather than looking at entrainment on target words, some studies investigate overall lexical entrainment. Overall lexical entrainment in individuals with ASD is typically measured during more unstructured naturalistic conversations, where the proportion of shared vocabulary between participants is measured (e.g. Stabile and Eigsti, 2022; Patel et al., 2022; Fusaroli et al., 2023). The majority of studies that measured lexical entrainment on a more global level rather than on target words, typically during less restricted conversations, report significant between-group differences, with individuals with ASD exhibiting lower degrees of lexical entrainment.

The present study aims to combine approaches taken in previous experiments: we will measure entrainment on target words, but will record more naturalistic, task-oriented conversations, with less predictable turn-taking than traditional studies in which entrainment on target words is measured. We hypothesise that a less structured and less predictable task will increase the cognitive load of participants, as they have to spend cognitive resources on the task, as well as on predicting turn-taking and other communicative and social processes. We hypothesise that the cognitive load during a semi-structured conversation is higher for people with ASD than NT people due to their differences in social processing. Since increased cognitive load is associated with reduced entrainment (Abel and Babel, 2017), we expect any between-group differences in entrainment to be more salient during semi-naturalistic conversation than in a highly structured one. In line with existing research, we expect to find less entrainment on the lexical level in our participants with ASD than in their NT peers.

3 Methods

3.1 Participants

For this experiment we collaborated with the Academic Research Center for Autism (ARCA) in Bratislava, Slovakia. With their help, we recruited two groups of children who were native Slovak speakers and had normal to corrected sight and hearing: one group of children with (suspected) ASD (diagnosis was later confirmed through standardised diagnostic testing) and one group of NT children who did not have suspected ASD or other developmental disorders.

In total, we recruited 67 children (14F, 62M), of whom 41 were diagnosed with ASD (7F, 34F) and 35 were NT (7F, 28M). The mean age of all recruited children was 9.21 (±1.86). For further details on the demographic information and various test scores of both groups, see Table 1. Note that we did not include data from each child in the analyses due to some technical issues with our audio recordings.

All children suspected of having ASD underwent a comprehensive diagnostic procedure, consisting of the Autism Diagnostic Observation Schedule (ADOS, Lord et al. (2008)) and the Autism Diagnostic Interview (Revised, ADI-R). Furthermore, the Woodcock-Johnson test was administered to assess the intelligence quotient (IQ) of each child,

Table 1: Summary of demographic information and test scores for both groups of participants.

	ASD		NT		t-test
	mean (std)	range	mean (std)	range	p
Age	9.10 (1.71)	6.14 - 12.30	9.34 (2.04)	6.18 - 12.97	>0.05
IQ	96.80 (16.81)	52 - 131	105.8 (14.84)	67 - 134	<0.05
BRIEF	67.32 (9.61)	47 - 85	58.77 (12.81)	36 - 83	<0.01
ToM	8.29 (2.14)	2 - 14	12.17 (1.74)	8 - 15	<0.001

while the Comic strip task (Cornish et al., 2010) was employed to measure Theory of Mind (ToM) abilities. Additionally, information regarding the executive functions (EF) of each child was gathered through the Behavior Rating Inventory of Executive Function (BRIEF) questionnaire (Gioia et al., 2000), which was completed by the parent(s)/caregiver(s) of the children. All materials used in the study were in Slovak and all tests were administered by a trained clinician at PhD-level.

3.1.1 Maps task

To elicit semi-naturalistic, task-oriented conversation, we used the Maps task. During this task, an experimenter and participant are both given maps that differ slightly (see Figure 1). One map contains a pre-drawn route, and the goal of the task is for the "instruction giver" to explain this route to the "instruction follower", so that the instruction follower can replicate the route on their own map as closely as possible.

We edited the original maps to change the original landmarks to different objects. Our maps contained two types of objects: control objects, which had one clearly preferred lexical label (e.g. *orech* (walnut) in Figure 1), and target objects which had both a "preferred" and "dispreferred" label (e.g. for the picture of the orange in in Figure 1, the preferred term was *pomaranč* (orange) and the dispreferred term was *mandarínka* (mandarin)). In the present experiment, we aimed to see whether children would entrain to the adult experimenter on dispreferred terms for target objects.

We selected our target objects based on an online norming study, in which we distributed a survey that contained coloured pictures and asked children to answer the following two questions: "What is the first word you would use to describe this picture?" and "What other word would you use to describe this picture?". Based on these answers, we selected our control objects (words that had one clearly preferred term and no common dispreferred terms) and target objects (words that had a preferred term and a commonly provided less-preferred or dispreferred term).

To minimise any discomfort or distress for the children with ASD, we decided that the experimenter who completed the Maps task with them would be somebody they were familiar with. The experimenter was the clinician who administered the other tests. The experimenter was thus aware of whether the child received an ASD diagnosis or not, which could introduce experimenter bias. To mitigate the influence of such potential bias, we provided the experimenter with training and detailed instructions on how to act prior to the task. Importantly, the experimenter was instructed to always use the dispreferred word, and as a reminder, her maps had written labels indicating with which word she should use to describe each object (see Figure 1).

The Maps task consisted of different trials, and different target objects were used in these different trials (see Table 2). The maps in the first and last "real" (i.e. non-practice) trial contained all 8 target items. The maps in trial 2 and 3 contained half of the target objects (the same 4 target objects each). This allows for the comparison of entrainment on dispreferred terms that were repeated more often and more recently to entrainment on dispreferred terms that were mentioned less often and a longer time ago. The maps used in trials 1 and 4 were counterbalanced, as were the maps used in trials 2 and 3.

In a typical Maps task, roles of participants (i.e. instruction giver or follower) switch between every trial. Since many children with ASD struggle with executive functioning, and such constant rule-switching might be challenging for them, we decided to only have one role switch in the Maps task: during the first few trials, the child was the instruction follower, and during the last few trials, the child was the instruction giver. Each part started with a practice trial (see Table 2), so we could ensure that the children understood the task and got used to their roles.

Figure 1: Example of maps used in the Maps task (not true to size). These specific maps were used for Trial 2, during which the experimenter is the instruction giver and only half of the target objects are on the map (see Table 2). The objects that are on both maps are target objects. Objects that are only present on one map are control objects that only have one preferred term. The target objects on the experimenter's map are labeled with their dispreferred terms as a reminder for her to use only the dispreferred terms.

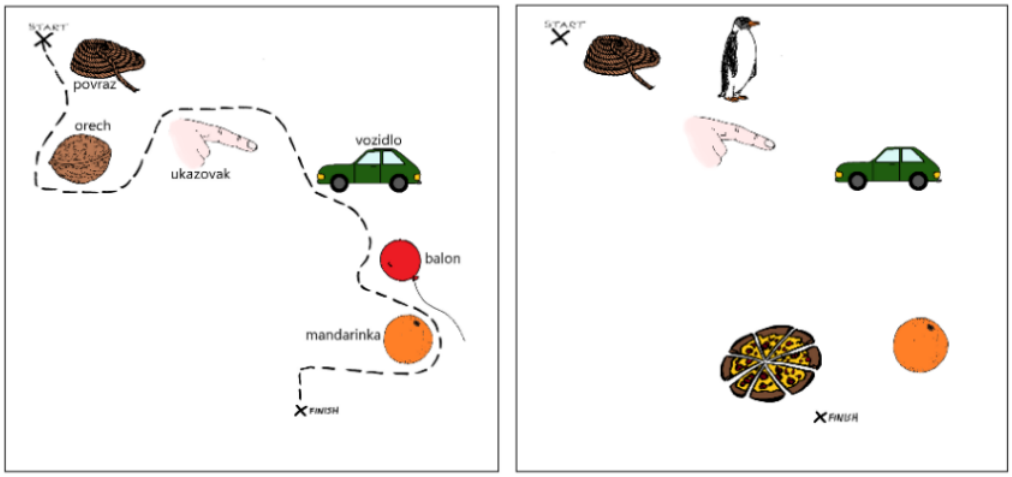


Table 2: Overview of Maps task trial structure.

Map	Instruction giver	Objects
Practice 1	Experimenter	Control only
Trial 1	Experimenter	All target objects
Trial 2	Experimenter	Half of the target objects
Practice 2	Child	Control only
Trial 3	Child	Half of the target objects
Trial 4	Child	All target objects

The Maps task was recorded with various microphones. All audio recordings were orthographically transcribed using Transcriber by experienced annotators who were native Slovak speakers. The subsequent transcriptions were transformed to Praat TextGrid format for further analysis.

3.2 Analysis

The Maps task elicits more naturalistic dialogue than many other tasks that have been used to research entrainment in children with ASD, and though this was an important goal of this project, it must be noted that it comes at a cost: the dialogue is unpredictable, and it is nearly impossible to predict how often each interlocutor will say a particular word. This applies to the child, but also to the experimenter. One could argue that if the experimenter says a dispreferred word more often, a child may be more likely to be primed to repeat it. For this reason, we included the number of times an experimenter said a particular word as a predictor of how many times the child used that same word.

To calculate lexical entrainment, we used the aforementioned TextGrids. One file was excluded because of technical issues, so 75 interactions were analysed. We used the Slovak *simplemma* lemmatiser to lemmatise all of each speaker's utterances and then counted how many times speakers said preferred and dispreferred words: per target word, we counted how many times the child used the preferred word and the dispreferred word, and how often the experimenter used the dispreferred word (since she had been instructed to always use the dispreferred word and never the preferred one). We then calculated the difference between the number of times the child said the preferred word and the dispreferred word, such that the number would be negative if the child dis-entrained, i.e. used the preferred word more often than the dispreferred. In the linear mixed effects model formula below, this value is represented as "diff child".

The linear mixed effects model formula we used to measure lexical entrainment was as follows, where "dispref exp" represents the number of times the experimenter used the dispreferred word:

diff child ∼ dispref exp + group + ToM + gender + trials + age + BRIEF + IQ + (1 | participant) + (1 | target item)

Group (i.e. ASD or NT), gender (M or F), ToM score, trials (i.e. whether the target stimuli was repeated in every trial or only in the first and last trials), age (in years), and BRIEF score are included as fixed effects, while participant and target stimulus are included as random effects. No interaction

effects were included as this seemed to lower the AIC of the model and thus indicated that the addition of these interaction terms did not lead to a better fit. The lmerTest R package (Kuznetsova et al., 2017), which provides p-values via Satterthwaite's degrees of freedom method, was used to assess significance of effects.

4 Results

The intercept of the model used to predict lexical entrainment, which corresponds to the values of group = ASD, trials = target item was repeated only in first and last trials, ToM score = 0, gender = F, number of times experimenter said dispreferred word = 0, age = 0, IQ = 0, and BRIEF score = 0, is presented in the first row of Table 3.

Table 3: Effects in the LMEM constructed for the lexical entrainment analysis. Significant effects are indicated with an asterisk.

effect	beta	std	t	df	p
intercept	1.29	1.78	0.73	75.44	0.470
dispref adult	0.19	0.03	5.40	533.08	<0.001*
group	-0.56	0.41	-1.37	68.66	0.174
ToM	0.06	0.08	0.73	68.68	0.471
gender	0.78	0.37	2.10	68.76	0.039*
trials	-0.57	0.79	-0.74	5.99	0.489
age	0.06	0.08	0.75	69.26	0.454
BRIEF	-0.03	0.01	-2.05	68.81	0.044*
IQ	-0.01	0.01	-0.92	68.77	0.359

Within this model, effects that were found to be non-significant and negative were group, target item repetition, and IQ score. Effects that were found to be non-significant but positive were those of ToM score, and age (see Table 3).

Several effects were found to be significant (see Table 3). The effect of the number of times the experimenter said the dispreferred word (dispref exp) was found to be significant and positive, suggesting that more repetitions of a term by the experimenter led to higher lexical entrainment on that term. Additionally, general BRIEF score was found to have a significant and negative effect on children's lexical entrainment, suggesting that increased issues with executive functioning (reflected in a higher BRIEF score) was associated with lower degrees of lexical entrainment. Finally, gender was found to be significant, suggesting that boys showed more lexical entrainment than girls (beta is positive and the intercept is for gender = F).

5 Discussion

Our two recruited groups did not differ significantly in age, though significant between-group differences existed for IQ, ToM, and BRIEF scores (see Table 1). The latter two are in line with existing research that suggests that children with ASD perform less well on ToM tests than NT children (e.g. Baron-Cohen, 2000; Baron-Cohen et al., 1985; Tager-Flusberg, 2007), and typically struggle more with executive functioning than NT children as well (see Demetriou et al. (2018) for a meta-analysis). We tried to match our groups as closely on possible on age and approximate IQ, but this is a difficult task. The group of NT children we recruited had a significantly higher mean IQ score than our group of children with ASD. Though this is not ideal for a between-group comparison, we added IQ as a fixed effect in our model and did not find that it predicted entrainment.

The present study aimed to assess lexical entrainment on target words, but during less constrained conversations than most existing studies. The results of the our analysis suggest that group (i.e. NT or ASD), ToM score, target item repetition, IQ score, and age do not significantly predict the degree to which a child entrained to the experimenter on a dispreferred term. On the contrary, the number of times the experimenter repeated a word and a child's BRIEF score both significantly predicted the child's lexical entrainment behaviour, such that more repetitions by the experimenter predicted higher entrainment, while a higher BRIEF score and thus more problems with executive functioning predicted lower lexical entrainment.

Our results are consistent with the majority of existing research that did not show between-group differences or significant effects of ToM ability in lexical entrainment on target words. Importantly, the only existing studies that indicated decreased entrainment in individuals with ASD during more unstructured, unpredictable dialogue assessed lexical entrainment in general, rather than on specific target words. In other words, these studies measured the proportion of shared vocabulary between interlocutors, rather than entrainment on specific lexical terms. It is possible that these different approaches to quantifying lexical entrainment may in reality measure two different conversational mechanisms or processes on different levels. Further research can elucidate whether measuring lexical entrainment in these different ways produces re-

sults that reflect the same underlying process, or whether lexical entrainment on a "global" versus more "local" scale perhaps rely on different mechanisms.

Based on the absence of significant effects regarding group membership or ToM scores, our findings do not appear to support the common ground/audience design account (Clark and Marshall, 1978; Clark and Murphy, 1982) of entrainment. Interestingly, our results suggest that certain social factors, such as gender, do significantly predict entrainment. This could be taken as support for the Communication Accommodation Theory by Giles et al. (1991).

Our results suggested that boys show more lexical entrainment than girls. A possible explanation for this is the observation that girls with ASD tend to use more compensatory strategies to "fly under the radar", or blend in in social settings. This behaviour is referred to as *camouflaging* or *masking* (Dean et al., 2017). It has been hypothesised that such strategies may also be used in language production (Parish-Morris et al., 2017) in interaction. It is possible that girls with ASD in this study were more likely to use the preferred word for an object because in everyday circumstances, this word would be used more commonly, and using a dispreferred word might make them stand out.

To see whether there was a difference in lexical entrainment between girls and boys between groups, we plotted the difference in preferred and dispreferred lexical item use by group and gender (see Figure 2). This figure shows that girls with ASD indeed show slightly less entrainment than boys and girls without ASD, though this difference is not significant. It is possible that camouflaging in girls with ASD and a general tendency towards social conformity that likely also exists in NT girls explains why girls show significantly less entrainment on dispreferred terms than boys.

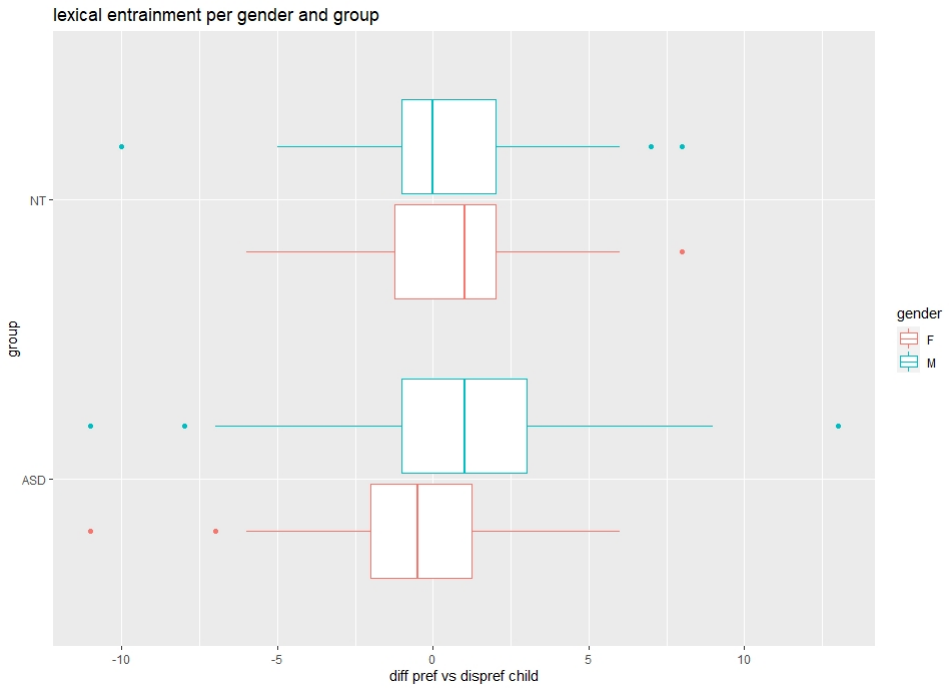
One could argue that our findings support Pickering and Garrod's interactive alignment hypothesis 2004; 2013: there is no difference in entrainment between groups, and no effect of ToM, suggesting that higher-order cognition is not required for entrainment. Additionally, the finding that the number of times the experimenter says a dispreferred word significantly predicts increased lexical entrainment in a child, supports the idea that priming underlies entrainment. However, Pickering and Garrod's theory 2004; 2013 does not explain why gender would affect entrainment, or why executive functioning significantly predicts the degree to which children lexically entrained to the experimenter.

Interestingly, few of the previous studies on entrainment in individuals with ASD included measures of executive functioning. Hopkins et al. (2016) investigated the effect of conflict inhibition on lexical entrainment and found no significant effects. Our results, which suggest that decreased executive functioning relates to lower lexical entrainment, are thus not in line with this previous study. A possible explanation of this is that in the Hopkins et al. (2016) study, a measure of one specific executive function, namely conflict inhibition, was included, and that this was measured with a specific test, whereas we asked participants' parents to fill out the BRIEF questionnaire as an indication of their executive functioning in everyday life. BRIEF scores may reflect a different set of executive functions than the test used by Hopkins et al. (2016).

An alternative explanation is that the experimental paradigm employed by Hopkins et al. (2016) was a game with single-utterance turns and a predictable conversation structure. The task used in the present study was more complicated and required active dialogue to complete. It is plausible that the increased cognitive load of our task and the accompanying dialogue required more of the children's cognitive resources than the game used by Hopkins et al. (2016). This could mean that there were fewer cognitive resources available for processes such as remembering that the experimenter used the dispreferred word and inhibiting the use of the preferred word, thus leading to decreased entrainment, especially in children who have more difficulties with tasks that require executive functioning skills. This is in line with existing research that has suggested that increased task demand and cognitive load leads to reduced entrainment (Abel and Babel, 2017).

The latter explanation is further supported by the study conducted by Stabile and Eigsti (2022), who also investigated lexical entrainment (on a global level) during a Maps task, and also measured executive functioning using the BRIEF questionnaire. While results of the study by Stabile and Eigsti (2022) did not reveal any significant associations between BRIEF score and lexical entrainment, results were marginally significant and in the same

Figure 2: Lexical entrainment, measured as the difference between the number of times a child used a preferred versus a dispreferred term, plotted by group and gender.



direction as the findings here: higher BRIEF score and thus more executive functioning difficulties were associated with lower lexical entrainment.

In other words, the results of the lexical entrainment analysis conducted in this experiment thus do not closely follow the predictions of any of the major theories of entrainment. Rather, the findings point towards a more nuanced and complex picture of lexical entrainment, in which various social and cognitive factors may influence the phenomenon.

6 Conclusion

The aim of this study was to investigate lexical entrainment on target words in children with and without ASD during a semi-naturalistic, task-oriented interaction that had less predictable turn-taking than previous studies. Results of our analysis suggest that some social factors such as age, and some (socio-)cognitive factors such as IQ and ToM score, do not significantly predict lexical entrainment. On the contrary, other social and cognitive factors, such as gender and executive functioning, do significantly predict lexical entrainment: girls show lower degrees of lexical entrainment on dispreferred terms than boys, and more executive functioning challenges in every day life are associated with decreased lexical entrainment. Moreover, the number of times an adult used a dispreferred word significantly predicted a child's entrainment on that dispreferred word. Taken together, the results of this study do not follow the predictions of any of the major theories of entrainment, suggesting that the phenomenon is complex and may be mediated by a number of different mechanisms and factors simultaneously.

Limitations

Dialogue was elicited using an experimental paradigm that does not have structured, predictable turn-taking. As with every decision made during a research process, this had advantages and disadvantages. A disadvantage of this decision was that we could not control the dialogue and thus could not control how often the experimenter used a dispreferred term. We tried to account for this by including it as a fixed effect in the LMEM we constructed, and found that this was indeed a significant predictor of entrainment. Importantly, we did not investigate the order in which interlocutors said dispreferred word: due to the unstructured nature of the conversation, it is possible that sometimes, a child referred to an object before the experimenter had a chance to refer to it by its dispreferred term. Future research may take a more qualitative approach, which could shed more light on the development of lexical entrainment during the dialogue.

Additionally, while our sample size was relatively large, we used a large statistical model and there is a chance our analysis was slightly underpowered. However, recruiting larger groups of children with ASD is extremely time- consuming and requires an incredible amount of resources, so this issue applies to many studies that aim to investigate behaviours of this population. Nonetheless, future studies may aim to implement different statistical tests to mitigate this issue.

Ethics Statement

ARCA's ethics board granted ethical approval for both the current experiment and the broader overarching research project in which this study was embedded. Prior to the experiment, informed consent was obtained from the parent(s) or caregiver(s) of the participants. To compensate for their time and participation, participants and their parent(s)/caregiver(s) received gift vouchers.

It is crucial to exercise caution when making assumptions about Theory of Mind (ToM) or social impairments in disorders such as Autism Spectrum Disorder (ASD). Traditionally, ASD has been associated with ToM impairments and inherent social deficits. However, recent empirical evidence challenges this assumption (e.g. Paynter et al., 2016; Gernsbacher and Yergeau, 2019). Instead of perceiving the communication, ToM, and social difficulties of individuals with ASD as their inherent deficits, it is proposed that these challenges arise due to "neurotype mismatches" occurring during interactions between individuals with ASD and neurotypical (NT) individuals. Individuals with ASD may not lack a theory of mind in general, but rather struggle to understand the mind of NT individuals specifically. Importantly, this perspective works in both directions, as NT individuals also seem to lack an understanding of the "autistic" mind (Sheppard et al., 2016; Heasman and Gillespie, 2018). This conceptualization is known as the "double empathy problem" (Milton, 2012), which is often advocated for by individuals with ASD. Given that most existing research on conversation coordination strategies of individuals with ASD has focused on interactions with a neurotype mismatch, it is crucial to consider this perspective.

Acknowledgements

We would like to thank all participants and their parent(s)/caregiver(s) for participating in this study. This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 859588, from the Slovak Granting Agency grant VEGA 2/0165/21, and the Slovak Research and Development Agency grant numbers APVV-21-0373 and APVV-20-0070.

References

Jennifer Abel and Molly Babel. 2017. Cognitive load reduces perceived linguistic convergence between dyads. *Language and Speech*, 60(3):479–502.

M.L. Allen, S. Haywood, G. Rajendran, and H. Branigan. 2011. Evidence for syntactic alignment in children with autism. *Developmental science*, 14(3):540–548.

S. Baron-Cohen. 2000. Theory of mind and autism: A review. *International review of research in mental retardation*, 23:169–184.

S. Baron-Cohen, A.M. Leslie, and U. Frith. 1985. Does the autistic child have a "theory of mind"? *Cognition*, 21(1):37–46.

G. Bossaert, H. Colpin, S.J. Pijl, and K. Petry. 2015. Quality of reciprocated friendships of students with special educational needs in mainstream seventh grade. *Exceptionality*, 23(1):54–72.

H.P. Branigan, A. Tosi, and K. Gillespie-Smith. 2016. Spontaneous lexical alignment in children with an autistic spectrum disorder and their typically developing peers. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 42(11).

H.H. Clark and C. Marshall. 1978. Reference diaries. In *Theoretical issues in natural language*, page 2.

H.H. Clark and G.L. Murphy. 1982. Audience design in meaning and reference. In *Advances in psychology*, volume 9, page 287–299. North-Holland.

K. Cornish, N. Rinehart, K. Gray, and P. Howlin. 2010. *Comic strip task*. Monash University Developmental Neuroscience and Genetic Disorders Laboratory and Monash University Centre for Developmental Psychiatry and Psychology, Melbourne.

Michelle Dean, Robin Harwood, and Connie Kasari. 2017. The art of camouflage: Gender differences in the social behaviors of girls and boys with autism spectrum disorder. *Autism*, 21(6):678–689.

E.A. Demetriou, A. Lampit, D.S. Quintana, S.L. Naismith, Y.J.C. Song, J.E. Pye, and A.J. Guastella. 2018. Autism spectrum disorders: a meta-analysis of executive function. *Molecular psychiatry*, 23(5):1198–1204.

J.W. Du Bois, R.P. Hobson, and J.A. Hobson. 2014. Dialogic resonance and intersubjective engagement in autism. *Cognitive Linguistics*, 25(3):411–441.

Riccardo Fusaroli, Ethan Weed, Roberta Rocca, Deborah Fein, and Letitia Naigles. 2023. repeat after me? both children with and without autism commonly align their language with that of their caregivers.

M.A. Gernsbacher and M. Yergeau. 2019. Empirical failures of the claim that autistic people lack a theory of mind. *Archives of scientific psychology*, 7(1):102.

H. Giles, N. Coupland, and J. Coupland. 1991. Accommodation theory: Communication, context, and consequence. In H. Giles, J. Coupland, and N. Coupland, editors, *Contexts of accommodation: Developments in applied sociolinguistics*, page 1–68. Cambridge University Press, Cambridge.

G.A. Gioia, P.K. Isquith, S.C. Guy, and L. Kenworthy. 2000. *Behavior rating inventory of executive function: BRIEF*. Psychological Assessment Resources, Odessa, FL.

B. Heasman and A. Gillespie. 2018. Perspective-taking is two-sided: Misunderstandings between people with asperger's syndrome and their family members. *Autism*, 22(6):740–750.

Z. Hopkins, N. Yuill, and B. Keller. 2016. Children with autism align syntax in natural conversation. *Applied Psycholinguistics*, 37(2):347–70.

A. Kuznetsova, P.B. Brockhoff, and R.H.B. Christensen. 2017. lmertest package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82(13):1–26.

C. Lord, M. Rutter, P.C. DiLavore, S. Risi, K. Gotham, and S.L. Bishop. 2008. *Autism diagnostic observation schedule (ADOS): Manual*. Western Psychological Services, Los Angeles.

D.E. Milton. 2012. On the ontological status of autism: the 'double empathy problem'. *Disability Society*, 27(6):883–887.

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):1–12.

Shivani P Patel, Jennifer Cole, Joseph CY Lau, Gabrielle Fragnito, and Molly Losh. 2022. Verbal entrainment in autism spectrum disorder and first-degree relatives. *Scientific reports*, 12(1):11496.

J.M. Paynter, D. Keen, and V.J. Rose. 2016. Systematic review documents limited empirical support for the practical application of the theory of mind model of asd. *Evidence-Based Communication Assessment and Intervention*, 10(3-4):131–139.

M.J. Pickering and S. Garrod. 2004. Toward a mechanistic psychology of dialogue. *Behavioral and brain sciences*, 27(2):169–190.

M.J. Pickering and S. Garrod. 2013. An integrated theory of language production and comprehension. *Behavioral and brain sciences*, 36(04):329–347.

E. Sheppard, D. Pillai, G.T.L. Wong, D. Ropar, and P. Mitchell. 2016. How easy is it to read the minds of people with autism spectrum disorder? *Journal of autism and developmental disorders*, 46(4):1247–1254.

K.E. Slocombe, I. Alvarez, H.P. Branigan, T. Jellema, H.G. Burnett, A. Fischer, and L. Levita. 2013. Linguistic alignment in adults with and without asperger's syndrome. *Journal of autism and developmental disorders*, 43(6):1423–1436.

Mackenzie Stabile and Inge-Marie Eigsti. 2022. Lexical alignment and communicative success in autism spectrum disorder. *Journal of Speech, Language, and Hearing Research*, 65(11):4300–4305.

H. Tager-Flusberg. 2007. Evaluating the theory-of-mind hypothesis of autism. *Current directions in psychological science*, 16(6):311–315.

A. Taheri, A. Perry, and P. Minnes. 2016. Examining the social participation of children and adolescents with intellectual disabilities and autism spectrum disorder in relation to peers. *Journal of Intellectual Disability Research*, 60(5):435–443.