Prosodic entrainment in individuals with autism spectrum disorder

Joanna Kruyt[*1, 2], Štefan Beňuš[1, 3]
[1]Institute of Informatics, Slovak Academy of Sciences in Bratislava, Slovakia
[2] Faculty of Informatics and Information Technology, Slovak Technical University in Bratislava, Slovakia
[3] Constantine the Philosopher University in Nitra, Slovakia

Abstract

Entrainment is the tendency of people to behave similarly during an interaction. It occurs on different levels of behaviour, including speech, and has been associated with pro-social behaviour and increased rapport. This review paper outlines the current understanding of linguistic entrainment, particularly at the speech level, in individuals with autism spectrum disorder (ASD), a disorder that is associated with social difficulties and unusual prosody. Aberrant entrainment patterns in individuals with ASD could thus contribute to both their perceived unusual prosody and their social difficulties. Studying the relationship between speech entrainment and ASD holds great potential for applied benefits in utilizing this knowledge for pre-screening or diagnosis, monitoring progress longitudinally, and intervention practices. Our findings suggest that research on entrainment in ASD is sparse and exploratory, and the ecological validity of experimental paradigms varies. Moreover, there is little consistency in methodology and results vary between studies, which highlights the need for standardized methods in entrainment research. A promising way to standardize methods, facilitate their use, and extend them to everyday clinical practice, is by implementing automatic methods for speech analysis and adhering to open-science principles.

Key words

autism spectrum disorders, entrainment, alignment, interpersonal synchrony, prosody

1. Introduction

During an interaction, people tend to repeat or mimic each other's gestures, movements, and ways of speaking (e.g. Chartrand and Bargh, 1999). As a result, the behaviour of two conversing interlocutors might become more similar over time. The tendency of people to behave more similarly in an interaction is often referred to as *entrainment*, which is the term that we will use in this paper, but other terms such as *convergence*, *synchrony*, *alignment* and *accommodation* are also commonly used. One form of behavioural entrainment that practically everyone is familiar with is yawning: this "contagious" behaviour is often mirrored and can spread around groups (Platek et al., 2003). Entrainment has also been observed in a large range of other modalities and behaviours, including posture sways (e.g. Shockley, Santana and Fowler, 2003), eye movements, head movements, and facial expressions (e.g. Louwerse, Dale, Bard and Jeuniaux, 2012), but also in physiological processes such as heartbeat (e.g. Feldman et al., 2011).

Perhaps the most well-researched domain in which entrainment occurs is language. It is this type of entrainment with which the present paper is concerned. Hence, in this paper, "entrainment" will denote

---

* Address for correspondence: Joanna Kruyt, Institute of Informatics (Slovak Academy of Sciences), Dúbravská cesta 9, Bratislava, Slovakia. E-mail: joanna.kruyt@savba.sk

linguistic entrainment unless otherwise specified. Entrainment has been observed at practically all levels of language, including syntax (e.g. Branigan, Pickering and Cleland, 2000), lexical choice (e.g. Brennan and Clark, 1996; Garrod and Anderson, 1987), prosody (e.g. Levitan et al., 2012; Natale, 1975; Webb, 1969), and dialogue acts (e.g. Mizukami et al., 2016). Entrainment is associated with smoother and more natural sounding conversation, as well as more effective and satisfying communication (e.g. Borrie, Lubold and Pon-Barry, 2015; Chartrand and Bargh, 1999; Fusaroli et al., 2012; Levitan et al., 2012; Nenkova, Gravano and Hirschberg, 2008; Reitter and Moore, 2014). Moreover, entrainment has been shown to correlate with feelings of closeness in interpersonal relationships (e.g. Lee et al., 2010; Pardo et al., 2012) and is associated with the building of rapport (e.g. Lubold and Pon-Barry, 2014). Additionally, entrainment in romantic relationships is positively associated with ratings of various measures of relationship quality such as cooperation and health of the relationship (Weidman, Breen, and Haydon, 2016) and positive attitude of partners (Lee et al., 2010). Entrainment has even been associated with attractiveness and likeability during speed dating events (Michalsky and Schoorman, 2017). In other words, entrainment plays a large role in interpersonal dynamics, social affiliation and conversation coordination.

1.1 Theoretical relevance

Despite its importance and relevance in many aspects of social interaction, several questions about entrainment remain unanswered. A heavily debated issue is whether entrainment occurs automatically (Pickering and Garrod, 2004, 2013), or is mediated by social factors (Giles, Coupland, and Coupland, 1991) or higher-order cognitive processes (Clark, 1996; Clark and Murphy, 1982) such as Theory of Mind (ToM), which refers to one's ability to comprehend and infer that others can have inner mental states that differ from one's own. Certain disorders, such as autism spectrum disorder (ASD) have traditionally been associated with ToM impairments (e.g. Baron-Cohen, 2000; Baron-Cohen et al., 1985; Tager-Flusberg, 2007). Studying entrainment in such populations may bring relevant findings bearing on the predictions of these theories, though caution must be taken regarding any conclusions about the role of ToM: while traditionally ASD has been associated with ToM impairments, recent empirical evidence suggests that this may be a questionable assumption (Paynter, Keen, and Rose, 2016; Gernsbacher and Yergeau, 2019). Rather than individual ToM impairments in individuals with ASD, difficulty in communications and "mind-reading" are the result of "neurotype mismatches" that occur when an individual who is not neurotypical, such as someone with ASD, interacts with a neurotypical (NT) individual. Individuals with ASD may not lack a theory of mind per se, but rather may struggle to understand the NT mind, specifically. Importantly, this works both ways: NT individuals lack a theory of "autistic" mind. This view is referred to as the "double empathy problem" (Milton, 2012) and is often advocated for by individuals with ASD. It is essential to keep this perspective in mind, especially considering the fact that all existing research into entrainment in individuals with ASD was conducted on interactions with a neurotype mismatch.

One of the most influential theories of entrainment was proposed by Pickering and Garrod (2004) in their seminal work on "interactive alignment". The idea at the core of the interactive alignment theory is that entrainment occurs automatically via priming: if speaker A hears speaker B say the word "couch", for example, this activates "couch" in speaker A's mental lexicon, and this recent activation makes it more likely that speaker A will use the word "couch" rather than the word "sofa" to describe a soft, comfortable multi-person seat. According to Pickering and Garrod (2004), this occurs due to the intricate link between production and perception of language: the perception of a specific linguistic form facilitates its production because perception and production rely on the same representations. These processes occur at all levels of language, including the prosodic, syntactic and conceptual levels. Pickering and Garrod (2004) further propose that entrainment "percolates" up through the different levels of language. For example, hearing the utterance "That is a good idea!" may prime someone to use the word "good". Subsequently, they may be more likely to say "That was a good movie" than other sentences with nearly identical meanings but different syntactic structures, such as "I liked that movie". In this way, lexical entrainment leads to entrainment at a higher level, namely at the level of syntax. Ultimately, this percolating of entrainment at various levels leads to what Pickering and Garrod call "alignment": shared conceptual representations of the dialogue. While Pickering and Garrod themselves do not explicitly discuss prosodic entrainment in their interactive alignment theory, it has been proposed that prosodic entrainment is the result of priming (e.g. Ward and Litman, 2007).

Contrary to the interactive alignment theory, other theories propose that entrainment is not automatic, but is mediated by social, cognitive or contextual factors. One of these theories is the Communication Accommodation Theory (CAT, Giles et al., 1991). According to CAT, people entrain to either emphasize or minimize the social distance between themselves and their conversational partner. This is comparable to the way body language has been shown to reflect interpersonal relationships or emotions: individuals who like each other tend to lean towards one another, thereby minimizing the distance between them. On the contrary, an individual who does not like their interlocutor typically leans away from them (e.g. Mehrabian, 1968; Trout and Rosenfeld, 1980). CAT states that entrainment is a similar phenomenon: mirroring someone, or entraining to them, minimizes social distance, while (perhaps purposefully) not entraining or even dis-entraining emphasizes this social distance (Giles et al., 1991). A prime example of this can be found in courtrooms where opposing parties use different words to emphasize their disagreement, for example the word "baby" versus "foetus" in an abortion trial (e.g. Clark, 1997; Danet, 1980). Indeed, research has shown that people tend to entrain less to individuals that are most different to them: for example, Menshikova, Kocharov and Kachkovskaia (2021) found greater entrainment between siblings than between strangers that differed in age or gender. Social distance thus appears to influence entrainment.

Although ToM has no direct role in entrainment according to CAT, it may affect it indirectly. ToM is essential for social cognition, which implies that impaired ToM may lead to impaired social cognition and thus perhaps also impaired social behaviour. Moreover, the neurotype mismatch that is present in communication between neurodivergent individuals (such as those with ASD) and NT individuals, may increase feelings of social distance and could thus in turn lead to different entrainment behaviours. This prediction is supported by a study conducted by Morrison et al. (2020), who found that individuals with ASD seemed to prefer having unstructured conversations with others who shared their diagnosis and neurotype, rather than with NT individuals. The authors suggest viewing the communication difficulties of individuals with ASD as "relational" impairments rather than a communication impairment of individuals with ASD, which is in line with the double empathy theory. Similarly, Crompton et al. (2020) found that neurotype-matched dyads reported higher levels of report in semi-structured interactions than dyads with a neurotype mismatch. These findings suggest that both interlocutors may experience more social distance in an interaction with a neurotype mismatch than when both interlocutors share a neurotype. Additionally, it is possible that individuals with ASD experience or perceive social distance in a different way than neurotypical individuals. Investigating entrainment behaviour in this group may thus shed light on how social distance is perceived and affects behaviour during interaction.

An alternative theory that assumes entrainment is not automatic is that of common ground and audience design. According to Clark and Marshall (1978), in order to communicate, interlocutors must share knowledge that must be mutually acknowledged as shared knowledge. Information that is mutually acknowledged to be shared is said to be "in the common ground" (Clark, 1996; Clark and Murphy, 1982). A concept closely related to the notion of common ground is "audience design" (e.g. Clark and Murphy, 1982) which refers to the way someone changes how they are speaking depending on who they are speaking to. In doing so, common ground is typically taken into account: "We have a good idea of the knowledge and beliefs they [the people we are talking to] share with us at the moment and what they are thinking of, and we design our utterances accordingly" (p. 287, Clark and Murphy, 1982). A prime example of this comes from a study by Bortfeld and Brennan (1997) where native and non-native speakers of English took part in a referential communication task in which participants had to describe and refer to objects or pictures. Once a non-native speaker described a rocking chair as "the chair in which I can shake my body" to a native speaker, the native speaker continued using the unusual description throughout the conversation rather than simply saying "the rocking chair" (p.132, Bortfeld and Brennan, 1997). This example illustrates how the native speaker designed their utterance for their non-native audience based on information in the common ground (in this case, the highly idiomatic expression). In this way, common ground and audience design can lead to entrainment.

In order to assess whether a piece of information is in the common ground, one must infer their interlocutor's knowledge and beliefs. This requires ToM. Subsequently, ToM is an essential component of audience design. Individuals with ASD appear to use information in the common ground differently to neurotypical individuals: Nadig, Seth and Sasson (2015) found that during a referential communication task, individuals with ASD and their interlocutors developed partner-specific terms to refer to objects, i.e. conceptual pacts, in a similar way to neurotypical participants. When participants

had to repeat the task with a new partner, neurotypical individuals repeated the process of establishing conceptual pacts, whereas individuals with ASD exhibited a tendency to re-use terms they had established with their old partner. It thus seems that these individuals may store or access information in the common ground differently, which may influence their audience design process and thus their utterances.

Although one may not expect prosodic entrainment as a result of audience design, it is important to remember that prosody is used to mark and decode shared or new knowledge in conversation, and thus common ground, through the distribution and prosodic characteristics of pitch accents and prosodic boundaries. Nonetheless, effects of audience design are more likely to be noticeable on other levels of language such as the lexical level. Moreover, considering the debate about ToM abilities and ASD, it could be argued that this theory is least likely to offer concrete predictions about prosodic entrainment in individuals with ASD.

In sum, investigating the patterns of prosodic entrainment in individuals with ASD might provide relevant knowledge for assessing the predictions of several leading theories of spoken communicative interaction.

1.2 Clinical relevance

Conclusive findings regarding the entrainment behaviour of individuals with ASD may not only be theoretically relevant, but could also be of clinical importance. Research has shown that individuals with ASD report more difficulties with long-term friendships and relationships and face more social exclusion than NT individuals (e.g. Bossaert et al., 2015; Taheri et al., 2016). Given the strong associations between entrainment and social affiliation, it is plausible that these social difficulties are partially caused or enhanced by aberrant entrainment patterns. If this is the case, focusing on entrainment as a therapeutic target in speech and language therapy may ameliorate the social difficulties that individuals with ASD experience when interacting with their NT peers. There is some evidence to support the notion of interventions targeting entrainment. For example, a therapy that focused on rhythmic entrainment of movement during a game (Srinivasan et al., 2005) reduced negative symptoms in individuals with ASD. Moreover, Koehne, Behrends, et al. (2016) found that adults with ASD who participated in a treatment focused on movement imitation and entrainment showed greater improvement in their ability to infer emotions than adults with ASD who participated in a different movement intervention that did not emphasize entrainment. Similarly, Ingersoll (2003) found that children with ASD who received treatment focused on imitation increased not only their imitation skills, but also their communication skills and joint attention, compared to children with ASD who did not receive such treatment. Despite these initial promising findings, further research is required to fully understand the potential of entrainment-based interventions for ASD, specifically in the domain of language.

A potentially interesting avenue of research regarding linguistic entrainment as a therapeutic target is the use of social robots or expressive avatars. Entrainment skills could be practised with social robots, who can model behaviour and provide feedback on an individual's performance (Diehl et al., 2012; Boucenna et al., 2014). Existing research suggests that individuals with ASD enjoy interacting with technology (Diehl et al., 2012; Boucenna et al., 2014) and that these interactions may even improve their social skills (Hopkins et al., 2011; Boucenna et al., 2014). Perhaps the spoken dialogue system implemented in social robots, avatars or other virtual agents could be designed for the purpose of speech and language therapy to encourage individuals with ASD to entrain more, with the goal of minimizing their social difficulties in real-world interactions.

Insight into entrainment behaviours of individuals with ASD may not only be useful in the development of potential interventions, but may also facilitate the diagnostic process. Early diagnosis of ASD is crucial as it allows for early intervention and can significantly improve quality of life for a child with ASD and their family (e.g. Elder et al., 2017; Fernell, Eriksson and Gillberg, 2013). However, getting a formal diagnosis is complicated: waiting lists for assessments are often long, and the process of testing the child and interviewing family members is time-consuming and requires certified experts. While it is clear that tests of linguistic entrainment cannot replace a comprehensive ASD assessment needed to evaluate and support an individual on the autism spectrum, any tools that can be used to screen children to identify those who might benefit from an official diagnostic assessment (Parish-Morris, Cieri, et al., 2016) or otherwise assist the diagnostic procedure could speed up the process.

Prior research supports the idea that children with ASD can be automatically distinguished from their NT peers based on their language and speech: several studies, including Parish-Morris, Liberman et al. (2016), Deng et al. (2017), and Lee et al. (2013) have shown that various machine-learning algorithms using features extracted from speech recordings can automatically identify children with ASD, NT children, and children with other (developmental) disorders at an above-chance level based on their acoustic-prosodic and lexical features. Parish-Morris et al. (2016) even provide a list of what they found to be the most and least "ASD-like" words, suggesting that there is a significant difference between the words used by both groups. Additionally, prosodic features such as speech rate and fundamental frequency (f0) facilitated the distinction between the three groups. In the future, including measures of entrainment in similar analyses could perhaps increase their accuracy. The inherent interactive nature of entrainment may facilitate the detection of individuals with social difficulties, such as individuals with ASD. Additionally, automatic detection of entrainment in individuals with ASD may be used for longitudinal monitoring of progress or assessing the effects of treatments and intervention, considering its strong associations with positive social measures. For all the aforementioned purposes, automatic speech analysis rather than (partially) manually conducted analyses requires less linguistic expertise and could save precious time and resources and may thus be favourable.

It must be noted that many tools to distinguish children with ASD from NT children based on language require the children to have reached a specific level of language development, and many children with ASD exhibit disordered or delayed language development. While prosodic entrainment may be determined on the basis of non-verbal vocalizations, this is not commonly done. Nonetheless, the development of screening tools may benefit verbal individuals from groups that are under-diagnosed with ASD, such as girls (e.g. Gould and Ashton-Smith, 2011; Dean, Harwood, and Kasari, 2017) or children of colour (e.g. Durkin et al., 2017).

1.3 Focus of the present paper

This paper aims to review and critically evaluate existing literature on linguistic entrainment in individuals with ASD. Research into entrainment on different levels of language will be reviewed, but a particular emphasis will be placed on prosodic entrainment. The speech and prosody of individuals with ASD is often perceived as odd or unusual (e.g. Nadig and Shaw, 2012; Filipe et al., 2014), which lends support to the notion that individuals with ASD may show unusual entrainment patterns, especially at the level of speech. Additionally, this review paper will focus on the ecological validity of existing research. Many (psycho)linguistic studies are conducted in laboratory environments, rely on highly controlled stimuli and involve predictable turn-taking. It is likely that participants in this setting thus have to rely less on conversation coordination strategies such as entrainment. Therefore, such highly controlled and predictable tasks are less likely to reveal differences in entrainment between individuals with and without ASD. To obtain results that more closely resemble everyday conversations, eliciting (semi-)naturalistic conversations from participants to investigate entrainment is imperative. Finally, we also discuss how the effective use of automatic methods for speech analysis and adherence to open-science principles might facilitate the standardization of methodology and ultimately the wider application of speech entrainment functionality in clinical practice and human-machine interactions.

2. Linguistic entrainment in individuals with autism spectrum disorder

2.1 Syntactic and lexical entrainment

Entrainment in individuals with ASD has been researched on several levels including syntax, lexical choice and acoustic-prosodic features. Research suggests that children with ASD show similar levels of syntactic entrainment to children without ASD, both in experimental settings (Allen et al., 2011; Slocombe et al., 2013) and more naturalistic conversations (Hopkins, Yuill and Keller, 2016). Similarly, research suggests individuals with and without ASD lexically entrain in similar ways, though all research involves highly structured tasks with predictable conversation. For example, in Branigan et al. (2016) and Hopkins et al. (2017) participants were asked to play a card game with highly predictable turn-taking and no need for multi-word utterances. Similarly, Slocombe et al. (2013) administered a card placing task in which participants were asked to produce short location descriptions following a predictable turn-taking pattern.

The fact that these studies relied on highly structured tasks begs the question of whether these experimental findings translate to more naturalistic conversation. Indeed, there is reason to believe that individuals with ASD entrain on syntax and lexical choice less in "real-world" conversations: even a commonly used screening tool to assess the social communication abilities of children who might have ASD, the Social Communication Questionnaire (SCQ; Rutter, Bailey and Lord, 2003), mentions "Lack of ability to imitate words and sentences" as a characteristic of ASD. Additional research may thus be required to confirm whether the experimental findings regarding lexical entrainment in individuals with ASD in highly structured tasks are also applicable to more natural and ecologically valid spoken interactions.

2.2 Prosodic entrainment

Although results from syntactic and lexical entrainment research suggest no significant differences in entrainment behaviours between individuals with and without ASD, findings from studies investigating acoustic-prosodic entrainment suggest otherwise. Wynn, Borrie and Sellers (2018) for example conducted an experiment in which participants were asked to watch videos of a woman asking participants to describe the picture she was holding. Some of the videos had been altered to increase or decrease the woman's speech rate. Participants were instructed to speak for 15 seconds after every video. Speech rate of participants was determined by counting the number of syllables in each response and dividing it by the total response time. Analysis revealed that the group of NT adult participants entrained on speech rate, i.e. increased their speech rate after watching a video with higher speech rate and decreased their speech rate after watching a video with slowed speech rate. This entrainment effect was not found for the group of adults with ASD, nor for the group of NT children, nor the children with ASD. The authors interpret the results as implying that "speech rate entrainment requires a subset of skills that are generally acquired during childhood and that may be impaired, or essentially not acquired, in individuals with ASD" (Wynn and Borrie, p. 969). In particular, they suggest that rhythmic skills are required for speech rate entrainment, and that these skills have not fully developed in younger children. Importantly, the experimental paradigm implemented by Wynn et al. (2018) involved entraining to audiovisual recordings rather than actual humans, which makes it difficult to draw conclusions on how these results translate to real-life, spoken interactions.

Hogstrom et al. (2018) studied the extent to which individuals with and without ASD entrained on a variety of prosodic-acoustic features to other humans, rather than recordings. Participants were asked to participate in a Maps task (Anderson et al., 1991). Both participants were given a map with landmarks. One participant's map had a pre-drawn route on it. The participant with this map was the "instruction giver" and had to give instructions to the other participant, the "instruction follower", whose goal was to replicate the pre-drawn route on their own map. The instruction giver's and follower's maps were similar, but not identical. Before and after the Maps task, participants were asked to read a list of sentences that contained the names of landmarks on the maps (e.g. "Number eight is the lighthouse", p. 1799). Hogstrom et al. (2018) then extracted the following features from these sentences: the spectral centre and duration of /s/ and /ʃ/ phonemes and the f0 and duration of the sentences. For each of these features, the mean absolute differences between speakers who completed the Maps task together were calculated, and the difference between the pre- and post-task recordings of the sentence lists were compared. If the mean absolute difference between speakers on a specific feature became smaller after completing the Maps task together, the participants had entrained on that particular feature. Interestingly, the only feature on which a significant difference was found between the group with and that without ASD was /s/ duration. Nonetheless, Hogstrom et al. (2018) conclude that this finding shows reduced entrainment in individuals with ASD. However, the robustness of this claim is questionable as a significant difference was found on only one of many features. Furthermore, entrainment was determined during monologic speech rather than during the interaction.

Lehnert-LeHouillier, Terrazas, and Sandoval (2020) also elicited task-oriented conversation from participants with and without ASD, and used this task-oriented dialogue rather than monologic speech to determine speech entrainment. Participants took part in a Diapix task (Baker and Hazan, 2011), which is essentially a collaborative "spot the differences" task. Each individual is given a picture. The pictures are similar but not identical, and participants must describe and discuss their images in order to find the differences between both pictures. To investigate entrainment on f0, the mean and interquartile range of f0 were extracted from the first and last third of every interaction. A geometric approach was taken

to quantify entrainment (see Lehnert-LeHouillier et al., 2020 for details). Results suggested a significant difference between groups in terms of f0 range as well as mean f0 entrainment: entrainment of on average 6 Hz was observed in conversations with NT participants, but in conversations with participants with ASD, dis-entrainment of on average 6 Hz was observed. There was also a significant difference between groups on mean f0 contribution: "on average, participants with ASD contributed 71% less to the convergence in mean f0 during the conversation compared to their NT peers" (Lehnert-LeHoullier et al., 2020 p. 8).

Furthermore, a significant interaction effect was found between age and group: results suggested that there was less entrainment in conversations involving older participants with ASD, while more entrainment was observed during conversations with younger participants with ASD. The opposite was found for the NT group: in conversations with older NT children, more entrainment was observed than during conversations with younger NT children. Importantly, this interaction effect was only found with regards to mean f0 entrainment, and not entrainment contribution. In other words, the difference originated from the confederate, rather than the participant. Lehnert-LeHoullier et al. (2020) state that perhaps the confederate was compensating for reduced entrainment in conversations with older participants with ASD, but that they did not do so in conversations with older NT children. However, it is difficult to know whether such adjustments were made consciously or not, as it is unclear whether the confederate was aware of the participants' diagnosis. Nonetheless, results from the study by Lehnert-LeHoullier et al. (2020) appear to indicate reduced entrainment in interactions with individuals with ASD as compared to NT individuals.

To date, the only study that studied entrainment in individuals with ASD during a semi-naturalistic conversation and within a smaller time-frame (i.e. not third or halves of a conversation) was done by Ochi et al. (2019), who analysed entrainment during the semi-structured interview of a diagnostic interview module (Lord et al., 2008). Ochi et al. (2019) used the time aligned moving average (TAMA) method (Kousidis et al., 2008) to assess entrainment. This method works by sliding a window across the speech signal at regular intervals and extracting features within this window to create a time series. The resulting time series are then cross-correlated. Ochi et al. (2019) extracted a large range of prosodic features as well as features relating to turn-taking. Important findings included a significant group difference on the correlation coefficient of "blockwise" mean intensity (i.e. the correlation coefficient between both speakers' mean intensity in a window) which was a measure of entrainment on intensity. Additionally, a support vector machine (SVM), a commonly used machine learning algorithm, was highly accurate in identifying individuals with ASD based on their speech features: it achieved an accuracy of 89.9 and an F-measure of 93.4 using the best combination of features which included mean and log standard deviation (SD) of turn-taking latency, SD of blockwise mean of intensity, and correlation coefficient of blockwise mean of log f0. In other words, individuals with ASD appeared to entrain less on intensity than NT participants, though this may have been related to the observation that individuals with ASD tended to speak at the same intensity and thus showed less variation in intensity. Using a range of features including a measure of f0 entrainment, a SVM could successfully identify individuals with ASD. This suggests that f0 entrainment differs between individuals with and without ASD, though it could also be related to the observation that more within-group variance was found on mean f0 in the ASD group as compared to the NT group. Taken together, the findings by Ochi et al. (2019) suggest that there may be reduced entrainment in individuals with ASD, though it is important to note that all participants in their study were male. While it is common for males to be overrepresented in ASD research (Watkins, Zimmerman and Poling, 2014), the absence of female participants in this study begs the question of whether its results are generalizable to the wider population with ASD. Nonetheless, the finding that a SVM could successfully detect individuals with ASD based on speech features and a measure of entrainment is promising. In general, research appears to suggest that there may be a pattern of reduced acoustic-prosodic entrainment in individuals with ASD, though only few studies analysed semi-naturalistic speech.

While Ochi et al. (2019) determined entrainment on a smaller timeframe than the previously discussed studies, their study still does not capture change in entrainment over time, i.e. "dynamic" entrainment. According to De Looze et al. (2014), fluctuations in entrainment over time may reflect changes in social behaviour, intentions or the speakers' involvement with each other. Considering the potential influence of social behaviour on entrainment and vice versa, capturing such dynamic changes might be essential to characterizing entrainment in general, but also specifically in individuals

with ASD. One way to determine entrainment dynamically is to use cross-recurrence quantification analysis (CRQA). This method has several benefits: besides capturing the dynamicity of entrainment, which may reflect changes in social behaviour, additional advantages of CRQA are the fact that it can be used to analyse naturalistic data semi-automatically, and that it can identify entrainment at different time lags (Fusaroli and Tylén, 2016; Borrie et al., 2019). A disadvantage of the method is that it requires the setting of several parameters that can critically alter the results of the analysis. Nonetheless, it is a promising tool for entrainment research more generally, as well as entrainment research in individuals with speech, language or communication difficulties, such as individuals with ASD.

Borrie et al. (2019, 2020) implemented CRQA in a way that is of particular relevance to entrainment research with clinical populations: they extracted a range of speech features from recordings of NT speech and conducted CRQA to quantify the extent of the behaviour that was shared between participants over time. Parameters for the CRQA were determined based on clinical assessments of how "in sync" interlocutors were. The inclusion of clinical assessments, and the fact that the method is automatic and suitable for naturalistic conversations, makes the CRQA-based method used by Borrie et al. (2020) an excellent option for future (clinical) research into entrainment in individuals with ASD. However, it is important to note that this particular implementation of CRQA requires clinical expertise and may thus be less appropriate for the automatic detection of individuals with ASD based solely on their speech features.

Table 1. Summary of studies investigating linguistic entrainment in individuals with ASD.

Level	Study	Findings	Setting
Syntactic	Allen et al. (2011)	ASD ≈ NT	Highly structured task
	Slocombe et al. (2013)	ASD ≈ NT	Highly structured task
	Hopkins et al. (2016)	ASD ≈ NT	Naturalistic dialogue
Lexical	Slocombe et al. (2013)	ASD ≈ NT	Highly structured task
	Branigan et al. (2016)	ASD ≈ NT	Highly structured task
	Hopkins et al. (2017)	ASD ≈ NT	Highly structured task
Acoustic-prosodic	Wynn et al. (2018)	ASD < NT	Monologic speech
	Hogstrom et al. (2018)	ASD < NT	Monologic speech
	Lehnert-LeHouillier et al. (2020)	ASD < NT	Semi-naturalistic dialogue
	Ochi et al. (2019)	ASD < NT	Semi-naturalistic dialogue

3. Discussion

The results of the existing literature review, summarized in Table 1, suggest that compared to individuals without ASD, individuals with ASD appear to show similar entrainment behaviours at the syntactic and lexical levels, but seem to exhibit less acoustic-prosodic entrainment. Interestingly, the majority of research into these topics was done using highly structured tasks. While the tasks involved interaction to some extent, often turn-taking was predictable and turns were short. For example, in the paradigm used by Branigan et al. (2016) and Hopkins et al. (2017), turns consisted of single-word utterances. In the task used by Slocombe et al. (2013), location descriptions were short. In both tasks, response options were constrained in terms of content and form. This could lead to higher levels of entrainment. Additionally, the paradigms require little actual interaction. Du Bois, Hobson and Hobson (2014) even refer to the interaction as "serial monologue" (p. 436) rather than a dialogue, highlighting how such structured tasks do not resemble true, naturalistic dialogue. Thus, the extent to which the results of these studies transfer to real-world conversations, in which turns are highly variable and commonly include incomplete linguistic structures, is a point of discussion. Future research could focus on more naturalistic conversations. However, it must be noted that not all speech acts in a naturalistic dialogue serve the same function, which means that dialogue act type may affect entrainment during the interaction. For example, Reichel, Mády and Cole (2018) found that people show a greater tendency to entrain during collaborative dialogue acts than during competitive ones. The way entrainment is affected by dialogue act type, social factors such as collaboration, and the interaction between the two requires further investigation, both in individuals with and without ASD.

Contrary to the other levels of language, existing research suggests that individuals with ASD exhibit different entrainment behaviours at an acoustic-prosodic level compared to individuals without ASD. This could be interpreted as not being in line with Pickering and Garrod's "interactive alignment" account of entrainment (2004, 2013), which states that entrainment occurs automatically through priming. This theory does not explain why acoustic-prosodic entrainment would differ in individuals with ASD compared to NT individuals: Pickering and Garrod state that "interlocutors do not model each other's mental states but simply align on each other's linguistic representations" (Pickering and Garrod, 2004, p. 180). An individual with a possible ToM impairment is thus unlikely to exhibit entrainment behaviours that differ from those produced by someone without such an impairment. Nonetheless, Pickering and Garrod mention that ToM may be used to correct for instances of "mis-alignment" or misunderstanding. Following the interactive alignment account, one would thus not expect an individual with ASD to entrain differently to a NT individual – except perhaps during misunderstandings.

It is important to note that Pickering and Garrod explicitly speak of linguistic representations. One could argue that acoustics and prosody serve more of a paralinguistic rather than purely linguistic function. This is a particularly crucial point when one realizes that existing research does not seem to find a difference between individuals with and without ASD in entrainment behaviours on any other linguistic levels, such as lexical choice and syntax (see Table 1). It is thus plausible that individuals with ASD do not have a linguistic entrainment deficit, but perhaps exhibit reduced entrainment on paralinguistic features such as voice quality, prosody and other non-verbal cues. This notion is supported by research that suggests that individuals have difficulty with synchrony and imitation in other modalities: for example, Marsh et al. (2013) found that children with ASD do not entrain to their parents on body swaying rhythm, whereas children without ASD do exhibit this tendency. Children with ASD also show reduced entrainment in pendulum swinging tasks (Fitzpatrick et al., 2018) and finger tapping tasks (Koehne, Hatri, et al., 2016). Patterns of reduced paralinguistic and non-verbal entrainment can also be observed in conversational settings: for example, Noel et al. (2017) showed that individuals with ASD show reduced entrainment on head and hand movements to their conversational partners as compared to NT individuals. Similarly, individuals with ASD do not entrain to their conversational partners on eye blinking (Nakano, Kato and Kitazawa, 2011). Future research could further explore this idea by investigating entrainment on paralinguistic features and other non-verbal cues. Additionally, as suggested by Wynn et al. (2018), the relationship between rhythmic skills and entrainment in this population could be investigated: a review of interpersonal synchrony in individuals with ASD illustrates that these individuals tend to show less entrainment than NT individuals in a range of behaviours including hand movements, head movements, clapping, pendulum swinging, marching and many other behaviours (McNaughton and Redcay, 2020). It is plausible that difficulties with rhythm underlie this lack of temporally synchronous behaviour, which may also influence entrainment on non-verbal communication and speech rate.

Alternatively, results of existing studies can be better explained by the audience design account, which postulates that during an interaction, individuals tailor their utterances to their conversational partner using perspective-taking and ToM skills (Clark, 1996; Clark and Murphy, 1982) that individuals with ASD are thought to be impaired in. However, this theory of mind impairment has been questioned Milton, 2012; Paynter, Keen and Rose, 2016; Gernsbacher and Yergeau, 2019) and recent evidence suggests the social difficulties experienced by individuals with ASD are the result of neurotype-mismatches rather than an inherent social (cognition) deficit (Morrison et al., 2020; Crompton et al., 2020). Thus, a perhaps more convincing interpretation of the findings is based on CAT (Giles et al., 1991). The results suggest that the social distance between an individual with ASD and a NT confederate may be perceived as greater than the social distance between two NT individuals, since they are of different neurotypes. This line of reasoning might apply to the observation in Lehnert-LeHoullier et al. (2020) that showed that the NT confederate entrained to the ASD interlocutors differently than to NT participants.

Although the results of the discussed studies suggest that individuals with ASD may entrain on prosodic features to a lesser extent than NT individuals, and sometimes may even disentrain, it is difficult to draw definitive conclusions. The tasks and analyses used in the studies differ greatly, as do their respective findings. For example, Wynn et al. (2018) did not measure entrainment to a human, but to video recordings. Hogstrom et al. (2018) measured entrainment based on word-list readings, rather than during interaction. Only Lehnert-LeHouillier et al. (2020) and Ochi et al. (2019) measured entrainment

during semi-naturalistic speech, though even those two studies differ significantly: the dialogue elicited by Lehnert-LeHouillier et al. (2020) was task-oriented while Ochi et al. (2019) collected speech samples during a semi-structured interview. Additionally, the two studies used vastly different analyses to quantify the degree of entrainment: Lehnert-LeHouillier et al. (2020) measured how similar two interlocutors became over the course of an entire conversation in absolute terms, whereas Ochi et al. (2019) looked at relative synchrony between two speakers in shorter timeframes. Such substantial differences in methods can complicate a direct comparison between studies. Additionally, it is possible that entrainment over large versus small timeframes occur due to different underlying mechanisms or serve different purposes in interaction. For this reason, it is essential that researchers specify which type of entrainment they are investigating and which timeframe is being considered (Levitan, 2018).

Various attempts have been made to provide frameworks for the classification of different types of entrainment, such as the framework proposed by Wynn and Borrie (2020). The fact that such frameworks exist indicates that the lack of standardized methodology and analysis is not specific to research investigating entrainment in ASD, but is characteristic of acoustic-prosodic entrainment research more generally (see e.g. Levitan, 2018, for discussion). An open-source entrainment analysis toolbox could facilitate the standardization of analyses in entrainment research, as well as make it more accessible for researchers whose backgrounds may be in other fields, including clinicians. Such a package already exists for the analysis of entrainment on different levels of language, such as lexical and syntactic entrainment (Duran, Paxton, and Fusaroli, 2019).

Furthermore, such a toolbox could facilitate the use of automatic methods. In addition to differences in various methodologies, the discussed studies relied partly on manual annotation: whether that be transcription (Lehnert-LeHouillier, Terrazas and Sandoval, 2020), extraction of inter-pausal units (Ochi et al., 2019) or counting of syllables to determine speech rate (Wynn, Borrie, and Sellers, 2018). The analysis process could be made more efficient and objective by carrying out these steps automatically. An open-source toolbox with computational methods that automatically extract markers of entrainment without the need for manual transcription would facilitate efficient and objective analysis, eliminate human error and ensure consistency in methodologies.

Entrainment research would also benefit from open data: not only would this facilitate replication of results; it would also allow for data to be re-analysed. Considering the dynamic and complex nature of entrainment (De Looze et al., 2014; Levitan, 2018) additional analyses could greatly improve our theoretical understanding of entrainment. There are myriad open-access corpora of NT individuals interacting, but similar resources are not available for the comparison of speech from NT individuals with that of individuals with ASD. Though this disparity is understandable considering the extra stringent personal data protection rules for data of clinical populations, such resources would be invaluable. Rather than sharing speech recordings, which may not always be ethically justifiable, perhaps researchers could share datasets of extracted acoustic-prosodic features and orthographic transcripts that have been anonymized. While this is by no means a perfect solution, it may be a small step towards increasing our understanding of entrainment in individuals with ASD, as well as facilitating research into other forms of speech analysis.

While adhering to open-science principles would be promising for acoustic-prosodic entrainment research more generally, specific suggestions can be made for further research into entrainment in ASD. The relationship between entrainment and perceived quality of conversation can be analysed to see whether this differs between NT individuals and those with ASD, as communication strategies may differ between the two groups: Morrison et al. (2020) found that individuals with ASD seemed to prefer having unstructured conversations with others who shared their diagnosis and neurotype, rather than with NT individuals. The authors suggest viewing the communication difficulties of individuals with ASD as "relational" impairments rather than a communication impairment of individuals with ASD. The role that entrainment plays in this relational perspective is worth investigating, as it is plausible that the different groups have preferences for different conversation styles or coordination strategies.

Based on Lehnert-LeHouillier, Terrazas and Sandoval (2020), the relationship between ASD symptom severity and entrainment behaviours should be further investigated. It could be researched whether acoustic-prosodic entrainment can provide a target for therapeutic intervention and also a means for gauging and monitoring long-term progress. Regarding potential intervention, either through human-human or human-machine spoken interactions individuals with ASD could be encouraged to entrain more, which may in turn enhance their social and conversational skills.

5. Conclusion

Research on linguistic entrainment in general and acoustic-prosodic entrainment in particular in individuals with ASD is sparse, and only recent studies employed naturalistic paradigms with higher ecological validity. Nonetheless, findings from existing studies suggest that individuals with ASD may exhibit patterns of reduced acoustic-prosodic entrainment or even show disentrainment, though future research is warranted to confirm and further characterize this possible entrainment deficit. In the future, open-science and computational, automatic methods could likely advance this field of research, and may have significant clinical implications, especially in the diagnostic process or as a potential therapeutic target.

Acknowledgment

References

Allen, M. L et al. 2011. Evidence for syntactic alignment in children with autism. *Developmental Science*, vol. 14, no. 3, pp. 540-548.

Anderson, A. H. et al. 1991. The HCRC map task corpus. *Language and Speech*, vol. *34*, no. 4, pp. 351-366.

Baker, R. and Hazan, V., 2011. DiapixUK: task materials for the elicitation of multiple spontaneous speech dialogs. *Behavior Research Methods*, vol. 43, no. 3, pp. 761-770.

Baron-Cohen, S., 2000. Theory of mind and autism: A review. *International Review of Research in Mental Retardation*, vol. *23*, pp. 169-184.

Baron-Cohen, S., Leslie, A. M. and Frith, U., 1985. Does the autistic child have a "theory of mind"?. *Cognition*, vol. 21, no. 1, pp. 37-46.

Borrie, S. A., Lubold, N. and Pon-Barry, H., 2015. Disordered speech disrupts conversational entrainment: a study of acoustic-prosodic entrainment and communicative success in populations with communication challenges. *Frontiers in Psychology*, no. 6, p. 1187.

Borrie, S. A. et al. 2019. Syncing up for a good conversation: A clinically meaningful methodology for capturing conversational entrainment in the speech domain. *Journal of Speech, Language, and Hearing Research*, vol. 62, no. 2, pp. 283-296.

Borrie, S. A. et al. 2020. Sync pending: Characterizing conversational entrainment in dysarthria using a multidimensional, clinically informed approach. *Journal of Speech, Language, and Hearing Research*, vol. 63, no. 1, pp. 83-94.

Bortfeld, H. and Brennan, S. E., 1997. Use and acquisition of idiomatic expressions in referring by native and non-native speakers. *Discourse Processes*, vol. 23, no. 2, pp. 119-147.

Bossaert, G. et al. 2015. Quality of reciprocated friendships of students with special educational needs in mainstream seventh grade. *Exceptionality*, vol. 23, no. 1, pp. 54-72.

Boucenna, S. et al. 2014. Interactive technologies for autistic children: A review. *Cognitive Computation*, vol. 6, no. 4, pp. 722-740.

Branigan, H. P., Pickering, M. J. and Cleland, A. A., 2000. Syntactic co-ordination in dialogue. *Cognition*, vol. 75, no. 2, B13-B25.

Brennan, S. E. and Clark, H. H., 1996. Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 22, no. 6, p. 1482.

Branigan, H. P., Tosi, A. and Gillespie-Smith, K., 2016. Spontaneous lexical alignment in children with an autistic spectrum disorder and their typically developing peers. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 42, no. 11, p. 1821.

Chartrand, T. L. and Bargh, J. A., 1999. The chameleon effect: the perception–behavior link and social interaction. *Journal of Personality and Social Psychology*, vol. 76, no. 6, p. 893.

Clark, H. H. and Marshall, C, 1978. Reference diaries. In *Proceedings of the 1978 Workshop on Theoretical Issues in Natural Language Processing*. pp. 57-63.

Clark, H. H., 1996. *Using language*. Cambridge: Cambridge University Press.

Clark, E. V., 1997. Conceptual perspective and lexical choice in acquisition. *Cognition*, vol. *64*, no. 1, pp. 1-37.

Clark, H. H. and Murphy, G. L., 1982. Audience design in meaning and reference. In *Advances in Psychology*, vol. 9, pp. 287-299.

Crompton, C. J. et al. 2020. Neurotype-matching, but not being autistic, influences self and observer ratings of interpersonal rapport. *Frontiers in Psychology*, vol. 11, p. 2961.

Danet, B., 1980. Baby or Fetus?: Language and the Construction of Reality in a Manslaughter Trial. *Semiotica La Haye* vol. 32, no. 3-4, pp. 187-219.

Dean, M., Harwood, R. and Kasari, C., 2017. The art of camouflage: Gender differences in the social behaviors of girls and boys with autism spectrum disorder. *Autism*, vol. 21, no. 6, pp. 678-689.

De Looze, C. et al. 2014. Investigating automatic measurements of prosodic accommodation and its dynamics in social interaction. *Speech Communication*, vol. 58, pp. 11-34.

Deng, J. et al. 2017. Speech-based diagnosis of autism spectrum condition by generative adversarial network representations. In *Proceedings of the 2017 International Conference on Digital Health*, pp. 53-57.

Diehl, J. J., et al. 2012. The clinical use of robots for individuals with autism spectrum disorders: A critical review. *Research in Autism Spectrum Disorders*, vol. 6, no. 1, pp. 249-262.

Du Bois, J. W., Hobson, R. P. and Hobson, J. A., 2014. Dialogic resonance and intersubjective engagement in autism. *Cognitive Linguistics*, vol. 25, no.3, pp. 411-441.

Durkin, M. S. et al. 2017. Autism spectrum disorder among US children (2002–2010): socioeconomic, racial, and ethnic disparities. *American Journal of Public Health*, vol. 107, no. 11, pp. 1818-1826.

Duran, N. D., Paxton, A. and Fusaroli, R., 2019. Align: Analyzing linguistic interactions with generalizable techniques—a python library. *Psychological Methods*, vol. 24, no. 4, p. 419.

Elder, J. H. et al. 2017. Clinical impact of early diagnosis of autism on the prognosis and parent–child relationships. *Psychology Research and Behavior Management*, vol. 10, pp. 283-292.

Feldman, R. et al. 2011. Mother and infant coordinate heart rhythms through episodes of interaction synchrony. *Infant Behavior and Development*, vol. 34, no. 4, pp. 569-577.

Fernell, E., Eriksson, M. A. and Gillberg, C., 2013. Early diagnosis of autism and impact on prognosis: a narrative review. *Clinical Epidemiology*, vol. 5, no. 33, pp. 33-43.

Filipe, M. G. et al. 2014. Atypical prosody in Asperger syndrome: Perceptual and acoustic measurements. *Journal of Autism and Developmental Disorders*, vol. 44, no. 8, pp. 1972-1981.

Fitzpatrick, P. et al. 2018. Relationship between theory of mind, emotion recognition, and social synchrony in adolescents with and without autism. *Frontiers in Psychology*, vol. 9, p. 1337.

Fusaroli, R. et al. 2012. Coming to terms: Quantifying the benefits of linguistic coordination. *Psychological Science*, vol. 23, no. 8, pp. 931-939.

Fusaroli, R. and Tylén, K., 2016. Investigating conversational dynamics: Interactive alignment, Interpersonal synergy, and collective task performance. *Cognitive Science*, vol. 40, no. 1, pp. 145-171.

Garrod, S. and Anderson, A., 1987. Saying what you mean in dialogue: A study in conceptual and semantic co-ordination. *Cognition*, vol. 27, no. 2, pp. 181-218.

Gernsbacher, M. A. and Yergeau, M., 2019. Empirical failures of the claim that autistic people lack a theory of mind. *Archives of Scientific Psychology*, vol. 7, no. 1, p. 102.

Gould, J. and Ashton-Smith, J., 2011. Missed diagnosis or misdiagnosis? Girls and women on the autism spectrum. *Good Autism Practice (GAP)*, vol. 12, no. 1, pp. 34-41.

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

Hogstrom, A. et al. 2018. Reduced phonemic convergence in autism spectrum disorder. Paper presented at *CogSci 2018: the 40th Annual Meeting of the Cognitive Science Society*. Madison, WI, United States, 25-28 July.

Hopkins, I. M. et al.2011. Avatar assistant: improving social skills in students with an ASD through a computer-based intervention. *Journal of Autism and Developmental Disorders*, vol. 41, no. 11, pp. 1543-1555.

Hopkins, Z., Yuill, N. and Keller, B., 2016. Children with autism align syntax in natural conversation. *Applied Psycholinguistics*, vol. *37*, no. 2, pp. 347-70.

Ingersoll, B. R., 2003. *Teaching children with autism to imitate using a naturalistic treatment approach: Effects on imitation, language, play, and social behaviors*. Doctoral dissertation: University of California, San Diego.

Koehne, S., et al. 2016. Fostering social cognition through an imitation-and synchronization-based dance/movement intervention in adults with autism spectrum disorder: A controlled proof-of-concept study. *Psychotherapy and Psychosomatics*, vol. 85, no. 1, pp. 27-35.

Koehne, S. et al. 2016. Perceived interpersonal synchrony increases empathy: insights from autism spectrum disorder. *Cognition*, vol. 146, pp. 8-15.

Kousidis, S. et al. 2008. Towards measuring continuous acoustic feature convergence in unconstrained spoken dialogues. Paper presented at the *Ninth Annual Conference of the International Speech Communication Association*. Brisbane, Australia, 22-26 September.

Lee, C. C., et al. 2010. Quantification of Prosodic Entrainment in Affective Spontaneous Spoken Interactions of Married Couples. Paper presented at the *Eleventh Annual Conference of the International Speech Communication Association*, Makuhari, Chiba, Japan, September 26-30.

Lee, H. Y., et al. 2013. Ensemble of machine learning and acoustic segment model techniques for speech emotion and autism spectrum disorders recognition. Paper presented at *Interspeech*, Lyon, France, August 25-29.

Lehnert-LeHouillier, H., Terrazas, S. and Sandoval, S., 2020. Prosodic entrainment in conversations of verbal children and teens on the autism spectrum. *Frontiers in Psychology*, vol. 11, p. 2718.

Levitan, R., 2018. Developing an integrated model of speech entrainment. In: *Proceeding of the Twenty-Ninth International Joint Conference on Artificial Intelligence (IJCAI-20) Eearly Career Track*, pp. 389-392.

Levitan, R., et al. 2012. Acoustic-prosodic entrainment and social behavior. In: *Proceedings of the 2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human language technologies*, pp. 11-19.

Levitan, R. and Hirschberg, J., 2011. Measuring acoustic-prosodic entrainment with respect to multiple levels and dimensions. Paper presented at the *Twelfth Annual Conference of the International Speech Communication Association*, Florence, Italy, August 27-31.

Lord, C., et al. 2008. *Autism diagnostic observation schedule (ADOS): Manual*. Los Angeles: Western Psychological Services.

Lubold, N. and Pon-Barry, H., 2014. Acoustic-prosodic entrainment and rapport in collaborative learning dialogues. In *Proceedings of the 2014 ACM Workshop on Multimodal Learning Analytics Workshop and Grand Challenge*, pp. 5-12.

Louwerse, M. M. et al. 2012. Behavior matching in multimodal communication is synchronized. *Cognitive science*, vol. 36, no. 8, pp. 1404-1426.

Marsh, K. L. et al. 2013. Autism and social disconnection in interpersonal rocking. *Frontiers in Integrative Neuroscience*, vol. 7, no. 4, pp. 1-8.

McNaughton, K. A. and Redcay, E., 2020. Interpersonal synchrony in Autism. *Current psychiatry reports*, vol. 22, no. 3, pp. 1-11.

Menshikova, A., Kocharov, A. and Kachkosvkaia, T., 2021. Lexical Entrainment and Intra-Speaker Variability in Cooperative Dialogues. In *Proceedings of Interspeech 2021*, pp. 1957-1961.

Mehrabian, A., 1968. Relationship of attitude to seated posture, orientation, and distance. *Journal of Personality and Social Psychology*, vol. 10, no. 1, p. 26.

Michalsky, J. and Schoormann, H., 2017. Pitch Convergence as an Effect of Perceived Attractiveness and Likability. In *Interspeech*, pp. 2253-2256.

Michalsky, J., Schoormann, H. and Niebuhr, O., 2018. Conversational quality is affected by and reflected in prosodic entrainment. in *Proceeding the 9th International Conference on Speech Prosody*, 2018, pp. 389-392.

Milton, D.E., 2012. On the ontological status of autism: the 'double empathy problem'. *Disability & Society*, vol. 27, no. 6, pp.883-887.

Mizukami, M. et al. 2016. Analyzing the effect of entrainment on dialogue acts. In *Proceedings of the 17th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, pp. 310-318.

Morrison, K. E. et al. 2020. Outcomes of real-world social interaction for autistic adults paired with autistic compared to typically developing partners. *Autism*, vol. 24, no. 5, pp. 1067-1080.

Nadig, A. and Shaw, H., 2012. Acoustic and perceptual measurement of expressive prosody in high-functioning autism: Increased pitch range and what it means to listeners. *Journal of Autism and Developmental Disorders*, vol. 42, no. 4, pp. 499-511.

Nadig, A., Seth, S. and Sasson, M., 2015. Global similarities and multifaceted differences in the production of partner-specific referential pacts by adults with autism spectrum disorders. *Frontiers in Psychology*, vol. 6, p. 1888.

Nakano, T., Kato, N, and Kitazawa, S., 2011. Lack of eyeblink entrainments in autism spectrum disorders. *Neuropsychologia*, vol. 49, no. 9, pp. 2784-2790.

Natale, M., 1975. Social desirability as related to convergence of temporal speech patterns. *Perceptual and Motor Skills*, vol. 40, no. 3, pp. 827-830.

Nenkova, A., Gravano, A. and Hirschberg, J., 2008. High frequency word entrainment in spoken dialogue. In *Proceedings of ACL-08: HLT, Short Papers*, pp. 169-172.

Ochi, K. et al. 2019. Quantification of speech and synchrony in the conversation of adults with autism spectrum disorder. *PloS One*, vol. 14, no. 12, p. e0225377.

Pardo, J. S. et al. 2012. Phonetic convergence in college roommates. *Journal of Phonetics*, vol. 40, no. 1, pp. 190-197.

Parish-Morris, J. et al. 2016. Exploring autism spectrum disorders using HLT. In *Proceedings of the conference. Association for Computational Linguistics. Meeting*, vol. 2016, p. 74.

Parish-Morris, J., et al. 2016. Building language resources for exploring autism spectrum disorders," in *Proceedings of LREC. International Conference on Language Resources & Evaluation*, p. 2100.

Paynter, J. M., Keen, D. and Rose, V. J., 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*, vol. 10, no. 3-4, pp. 131-139.

Pickering, M. J. and Garrod, S., 2004. Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, vol. *27*, no. 2, pp. 169-190.

Pickering, M. J. and Garrod, S., 2013. An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*, vol. 36, no. 4, pp. 329-347.

Platek, S. M. et al. 2003. Contagious yawning: the role of self-awareness and mental state attribution. *Cognitive Brain Research*, vol. 17, no. 2, pp. 223-227.

Reichel, U. D., Mády, K. and Cole, J., 2018. Prosodic entrainment in dialog acts. *arXiv preprint arXiv:1810.12646*.

Reitter, D. and Moore, J. D., 2014). Alignment and task success in spoken dialogue. *Journal of Memory and Language*, vol. 76, pp. 29-46.

Rutter, M., Bailey, A. and Lord, C., 2003. *The social communication questionnaire: Manual*. Los Angeles: Western Psychological Services.

Shockley, K., Santana, M. V. and Fowler, C. A., 2003. Mutual interpersonal postural constraints are involved in cooperative conversation. *Journal of Experimental Psychology: Human Perception and Performance*, vol. 29, no. 2, p. 326.

Slocombe, K. E. et al. 2013. Linguistic alignment in adults with and without Asperger's syndrome. *Journal of Autism and Developmental Disorders*, vol. 43, no. 6, pp. 1423-1436.

Srinivasan, S. M. et al. 2015. A comparison of the effects of rhythm and robotic interventions on repetitive behaviors and affective states of children with Autism Spectrum Disorder (ASD). *Research in Autism Spectrum Disorders 18*, pp. 51-63.

Tager-Flusberg, H., 2007. Evaluating the theory-of-mind hypothesis of autism. *Current Directions in Psychological Science*, vol. 16, no. 6, pp. 311-315.

Taheri, A. Perry, A. and Minnes, P., 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*, vol. 60, no. 5, pp. 435-443.

Trout, D. L. and Rosenfeld, H. M., 1980. The effect of postural lean and body congruence on the judgment of psychotherapeutic rapport. *Journal of Nonverbal Behavior*, vol. 4, no. 3, pp. 176-190.

Ward, A. and Litman, D. J., 2007. Automatically measuring lexical and acoustic/prosodic convergence in tutorial dialog corpora. Paper presented at the *SLaTE Workshop on Speech and Language Technology in Education*, Farmington, PA, USA, October 1-3.

Watkins, E. E., Zimmermann, Z. J. and Poling, A., 2014. The gender of participants in published research involving people with autism spectrum disorders. *Research in Autism Spectrum Disorders*, vol. *8*, no. 2, pp. 143-146.

Webb, J. T., 1969. Subject speech rates as a function of interviewer behaviour. *Language and Speech*, vol. 12, no. 1, pp. 54-67.

Weidman, S., Breen, M. and Haydon, K. C., 2016. Prosodic speech entrainment in romantic relationships. In *Proceedings of Speech Prosody*, pp. 508-512.

Wynn, C. and Borrie, S. A., 2020. Classifying Conversational Entrainment of Speech Behavior: An Updated Framework and Review. Uploaded on PsychArxiv on September 28. Available at https://doi.org/10.31234/osf.io/nrj2z.

Wynn, C. J., Borrie, S. A. and Sellers, T. P., 2018. Speech rate entrainment in children and adults with and without autism spectrum disorder. *American Journal of Speech-Language Pathology*, vol. 27, no. 3, pp. 965-974.