HHS Public Access

Coarticulation facilitates lexical processing for toddlers with autism

Ron Pomper[1,2], **Susan Ellis Weismer**[2,3], **Jenny Saffran**[1,2], **Jan Edwards**[4]

[1]Department of Psychology, University of Wisconsin-Madison, Brogden Hall 1202, West Johnson Street, Madison, WI 53706, USA

[2]Waisman Center, University of Wisconsin-Madison, 1500 Highland Avenue, Madison, WI 53705, USA

[3]Department of Communication Sciences and Disorders, University of Wisconsin-Madison, Goodnight Hall, 1975 Willow Drive, Madison, WI 53706, USA

[4]Department of Hearing and Speech Sciences, University of Maryland, Samuel J. LeFrak Hall, 7521 Preinkert Drive, College Park, MD 20742, USA

Abstract

Many children with autism spectrum disorder (ASD) are delayed in learning language. The mechanisms underlying these delays are not well understood but may involve differences in how children process language. In the current experiment, we compared how 3- to 4-year-old children with ASD (n=58) and 2- to 3-year-old children who are typically developing (TD, n=44) use phonological information to incrementally process speech. Children saw pictures of objects displayed on a screen and heard sentences labeling one of the objects (e.g., *Find the ball*). For some sentences, the determiner *the* contained coarticulatory information about the onset of the target word. For other sentences, the determiner *the* did not contain any coarticulatory information. Children were faster to fixate the target object for sentences with vs. without coarticulation. This effect of coarticulation was the same for children with ASD compared to their TD peers. When controlling for group differences in receptive language ability, the effect of coarticulation was stronger for children with ASD compared to their TD peers. These results suggest that phonological processing is an area of relative strength for children with ASD.

1. Introduction

Delays in learning language are common for many children with autism spectrum disorder (ASD). These delays extend across different aspects of language (i.e., semantics, syntax, morphology, phonology), vary extensively between children, and are associated with long-term outcomes (Kjelgaard & Tager-Flusberg, 2001; Pickles, Anderson, & Lord, 2014; Tager-Flusberg & Kasari, 2013). Although there is an expansive body of research examining language outcomes for children with ASD, considerably less is known about how children with ASD process language (Arunachalam & Luyster, 2016). Understanding how children process language is important because it may illuminate the mechanisms that lead to differences in language outcomes.

An important feature of language processing is that it occurs incrementally. Spoken language unfolds over time. Listeners incrementally process speech, anticipating the next words before they are heard. From a very young age, children who are typically developing (TD) incrementally process speech, identifying and looking to objects before they are labelled. Children use many different types of information to incrementally process speech, including semantics (e.g., *eat the cake*; Borovsky, Elman & Fernald, 2012; Fernald et al., 2008; Mani & Huettig, 2012), morphosyntax (e.g., Lew-Williams & Fernald, 2007; Lukyanenko & Fisher, 2016), pragmatics (e.g., Borovsky & Creel, 2014; Kidd, White, & Aslin, 2011), and phonology (e.g., Swingley, Pinto, & Fernald; 1999; Mahr, McMillan, Saffran, Ellis Weismer, & Edwards, 2015).

Many theories suggest that prediction is integral to both language processing and language learning (Chang, Dell, & Bock, 2006; Christiansen & Chater, 2016; Dell & Chang, 2013; Elman, 1990; Pickering & Garrod, 2013). Indeed, children who are better at predicting when processing language typically have larger vocabularies (Borovsky et al., 2012; Borovsky & Creel, 2014; Lew-Williams & Fernald, 2007; Mani & Huettig, 2012). Moreover, children who are better at predicting and revising incorrect predictions when processing language tend to learn novel words best (Reuter, Borovsky, & Lew-Williams, 2019). Together, this research demonstrates that prediction plays an important role in speech processing for TD children. Several experiments have compellingly demonstrated that children with ASD also use semantic information to incrementally process speech (Bavin, Kidd, Prendergast, & Baker, 2016; Brock, Norbury, Einav, & Nation, 2008; Hahn, Snedecker, & Rabagliati, 2015). Moreover, like their TD peers, children with ASD who are better at incremental language processing tasks typically have larger vocabularies (Venker, Edwards, Saffran, & Ellis Weismer, 2019). These findings demonstrate that children with ASD are able to rapidly generate predictions when processing speech.

In light of these findings, the research focus should shift from questioning whether children with ASD can incrementally process speech to instead examining under what circumstances they are able to do so. Studies to date have focused solely on the use of semantic information

for incremental speech processing in this group of children. Phonological information, however, may be particularly useful for children with ASD engaged in incremental speech processing. Prominent theories of autism propose that individuals with ASD have enhanced perceptual processing or an attentional style that is biased towards processing local information (Happé & Booth, 2008; Happé & Frith, 2006; Mottron & Burack, 2001). Several experiments have found that children with ASD are better at encoding phonological information embedded in words than their TD peers who are matched in verbal ability (Henderson, Powell, Gaskel, & Norbury, 2014; Nadig & Mulligan, 2017; Norbury, Griffiths, & Nation, 2010). Therefore, children with ASD may excel at incrementally processing speech using phonological information that is embedded in words.

In the current experiment, we used a looking-while-listening (LWL) paradigm to test whether children with ASD exploit phonological information during incremental speech processing. Specifically, we examined whether they are sensitive to coarticulation, the process by which individual sounds influence adjacent sounds in fluent speech (e.g., Daniloff & Hammarberg, 1973; Elman & McClelland, 1986). Coarticulation is ubiquitous in running speech and TD adult listeners take advantage of coarticulatory cues in both speech perception and word recognition (e.g., Gow & McMurray, 2007; Mattys, White, & Mellhorn, 2005; Salverda, Kleinschmidt, & Tannehaus, 2014). Recent research suggests that TD children as young as 18 months also use coarticulatory cues for spoken word recognition. Mahr and colleagues (2015) found that 18- to 24-month-olds were faster to fixate a target object when it was labelled using a sentence (e.g., *Find the ball*) where the preceding determiner *(the)* contained coarticulatory information about the onset of the target word (*ball*) compared to when coarticulatory cues were removed. We predicted that children with ASD would also use coarticulatory cues during incremental speech processing. Moreover, based on previous research suggesting that attention to phonological information is a relative strength for children with ASD, we predicted that children with ASD would be *more* sensitive to coarticulation than TD peers who were matched in receptive language ability.

2. Method

2.1 Participants

The final sample included 58 children with ASD (37–50 months; 15 females) and 44 younger TD children (18–36 months; 28 females). This age difference was intentional, because our goal was to match children in language ability, rather than chronological age (see Table 1). By matching children for language ability, we can attribute any differences between groups in speech processing to fundamental differences in language development, rather than developmental delays. This matching is particularly important because past research has shown that the extent to which children incrementally process speech is associated with individual differences in their receptive language ability (Borovksy & Creel, 2014; Lew-Williams & Fernald, 2007; Mani & Huettig, 2012; Venker et al., 2019).

All children in the ASD group met ASD criteria assessed using the Autism Diagnostic Observation Schedule, 2nd Edition (ADOS-2; Lord et al., 2012) and the Autism Diagnostic Interview, Revised (ADI-R; Rutter et al., 2003). Children in the

TD group were excluded if they were at an increased risk for ASD or there were concerns about developmental delay (Rutter, LeCouteur, & Lord, 2003). All parents were monolingual English speakers. More detailed information about participant exclusions is included in the Supplementary Materials.

2.2 Procedure

Participation involved two visits that were scheduled no more than 3 weeks apart. Each visit lasted approximately 1 hour for TD children and 2.5 hours for children with ASD. All children completed multiple LWL (Fernald et al., 2008) tasks that assessed different aspects of online language comprehension. We report the results for the coarticulation task here; results for the other tasks are reported elsewhere (Pomper, Ellis Weismer, Saffran, & Edwards, 2019; Venker, Haebig, Edwards, Saffran, & Ellis Weismer, 2016; Venker, Pomper, Mahr, Edwards, Saffran, & Ellis Weismer, 2019). In addition to the online measures of language comprehension, children and parents completed a set of standardized assessments and questionnaires.

2.3 Standardized Assessments

Children completed the Preschool Language Scales, 5th edition (PLS-5; Zimmerman et al., 2011) and the Mullen Scales of Early Learning (Mullen 1995). Children's receptive language ability was quantified using their raw score on the Auditory Comprehension scale (PLS-5) and their nonverbal cognitive ability was quantified using their raw score on the Visual Reception scale (Mullen). We used raw scores (rather than standard scores) to control for overall differences in verbal or nonverbal cognitive ability (rather than age-adjusted differences) between the groups (see section 2.5).

One child in the ASD group did not complete the PLS-5. Out of the 57 children with ASD who did complete the task, 41 displayed clinical language delays with a standard score below 81 on the PLS-5 (i.e., scores at least −1.25 SD below the mean). Two children in the TD group did not complete the PLS-5. Of the 42 children with TD who did complete the task, none displayed clinical language delays.

Despite our attempts to minimize group differences in receptive language ability by recruiting younger children into the TD group and older children into the ASD group, there were still significant group differences. Children in the ASD group had significantly lower raw scores on the PLS-5 [$Mean$=25.4, SD=11.9] than children in the TD group [$Mean$=33.7, SD=7.96], b=8.28, t=3.9, p<.001 (see Table 1). Children's raw scores on the Mullen, however, did not significantly differ between groups, b=2.6, t=1.6, p=.11. Therefore, we included children's raw score on the PLS-5 as a covariate in our analyses to control for group differences in receptive language ability (see section 2.5). We prefer this approach; the alternative approach of identifying subsamples of children from each group who were matched in PLS-5 leads to the exclusion of many children from the full sample and substantially reduces statistical power. For full transparency, we have included the matched subsample analyses in Supplementary Materials.

2.4 Experimental Task

The effect of coarticulatory cues on children's incremental speech processing was measured using a modified version of a paradigm from Mahr et al. (2015). On each trial, children were shown images of two objects that were displayed in silence for 1.5 seconds. Children then heard a sentence labelling one of the objects. The target words were presented in carrier phrases (e.g., *Find the ___*, *See the ___*). We manipulated whether the determiner *the* provided coarticulatory cues. On **Facilitating** trials, the determiner contained coarticulatory cues for the onset of the target word. On **Neutral** trials, the determiner did not contain coarticulatory cues for the onset of the target word. Using this paradigm, we measured the effect of coarticulation on speech processing by comparing children's word recognition accuracy (i.e., their accuracy in fixating the target image after it was labeled) on Facilitating and Neutral trials.

2.4.1 Materials and Stimuli—A female native English speaker with a local accent recorded multiple tokens for each sentence. Tokens were selected to have similar intonation contours and were edited using Praat to normalize intensity (RMS amplitude) and duration across items. For each carrier phrase, the determiner *the* was cross-spliced from another sentence. Determiners for Facilitating trials (*the*^{b}, *the*^{d}, or *the*^{g}) came from the phrases *find the ball*, *find the dog,* and *find the grapes*. Determiners for Neutral trials (*the*^{ə}) came from the phrase *find the hut*.

Images were color photographs of familiar objects. Each target word occurred on multiple trials, so we selected multiple images for each object to help maintain children's attention. Images were edited to match objects in size and placed on a gray background. We selected 12 objects whose labels were familiar to children in this age range, yoked into 6 pairs: *duck–ball*, *bed–door*, *drum–book*, *boots–grapes*, *grass–bowl*, and *dog–bus*. Pairings were chosen so that both target words had phonologically distinct onsets, belonged to different semantic categories, and were approximately matched in salience. Each object occurred equally often as the target and as the distractor, on the left and right side of the screen, and in each condition. There were a total of 48 trials, half with Facilitating coarticulatory cues and half with Neutral coarticulatory cues. Trials were divided equally into two blocks that were administered during separate visits.

2.4.2 Data collection, coding, and cleaning—Children were seated approximately 2 feet away from a 55-inch, wall-mounted TV. Sounds were played from a speaker mounted underneath the TV. Children sat either in their caregiver's lap or on their own with their caregiver standing behind them. Children's eye gaze was recorded using a video camera mounted underneath the TV.

For each video, trained coders indicated on each frame (33 ms) whether children were looking at the left image, right image, or neither image. To measure reliability, 20% of children in each group were coded independently by two coders. Trials that were not initially comparable (29.1% of trials for the ASD group and 19.0% of trials for the TD group) were discussed and coded by consensus. The proportion of all frames on which coders agreed on fixation location was 97.6% for the ASD group and 98.3% for the TD group. The mean

proportion of shifts in fixation location on which coders agreed within one frame was 93.3% for the ASD group and 95.4% for the TD group.

Before analyzing the data, we removed any trial where the child was inattentive (i.e., *not* looking at either of the images that were displayed on the screen for more than half of the critical window 300 to 900 ms after the onset of the target word). These trials were excluded because children did not contribute adequate data. Out of the possible 24 trials, children in the ASD group contributed on average 16.6 trials (SD=5.2) in each condition and children in the TD group contributed on average 19.8 trials (SD=3.8) in each condition. This difference between groups was statistically significant (b=3.3, p<.001) and was expected based on prior research (Ellis Weismer et al., 2016). The number of trials did not differ between conditions and the difference between conditions was not moderated by group (p's > 0.46).

2.5 Data Analysis

The dependent variable was the proportion of trials on which children were fixating the target image out of the trials they were fixating the distractor image. This proportion was calculated for each frame (every 33 ms) and transformed to weighted empirical logits (Barr, 2008). We used growth curve analysis (GCA) to model how children's probability in fixating the target image changed over time (Mirman, 2014). To address concerns that GCA may be flawed (Cho, Brown-Schmidt, & Lee, 2018; Huang & Snedeker, 2020), we validated our analyses, demonstrating that our GCAs were not anti-conservative and that the pattern of results was replicated using cluster-based permutation analyses (see Supplementary Materials).

Traditionally, word recognition accuracy is measured during a critical window 300 to 1,800 ms after the onset of the target word (Fernald et al., 2008). Fixations that occur before this window cannot be in response to the target word, because it takes children approximately 300 ms to program an eye movement. Similarly, fixations after this window are unlikely to be stimulus-driven, because children's attention wanes over time. In the current experiment, we set our critical window to be 300 to 900 ms after the onset of the target word. This shorter window was chosen for several reasons. First, coarticulation should have an early, but not a late, effect of word recognition accuracy. Consistent with prior work, we expected that coarticulation would facilitate children's ability to identify the referent, but not affect their ultimate, peak accuracy in fixating the referent (Mahr et al., 2015). Second, visual inspection of the raw data (see Figure 1) confirms that after 900 ms children's accuracy in fixating the target image begins to asymptote (for the ASD group) and differences between trials with and without coarticulation disappear. Finally, cluster-based permutation analyses (see Supplementary Materials) confirm that the effect of coarticulation occurs only during our critical window.

Children's weighted empirical logits were regressed on orthogonal time terms (intercept, linear, quadratic, and cubic), condition (contrast coded as −0.5 for Facilitating and 0.5 for Neutral), group (contrast coded as −0.5 for ASD and 0.5 for TD). Due to the significant group differences in receptive language ability (see section 2.3), we fit a second model that included children's raw score on the Auditory Comprehension scale of the PLS-5 (mean-

centered). A full list of the fixed effects is included in Tables 2 and 3 and model specifications are included in the Supplementary Materials.

Each orthogonal time term quantifies different geometric properties for how children's accuracy in fixating the target image changes throughout the critical window. The intercept quantifies the overall area under the curve, which is the average accuracy across the entire window. Linear time (ot1) quantifies the slope of the line, which is the monotonic increase in accuracy per unit of time (every 33ms). Quadratic time (ot2) quantifies the change in the slope of the line over time, which captures the degree to which increases in accuracy accelerate towards the end of the window. Cubic time (ot3) quantifies changes in the slope of the line around the tails, which captures asymptotes in children's accuracy at the beginning or the end of the critical window.

All models were fit using Maximum Likelihood estimation and included participant and participant-by-condition random effects. Analyses were performed in RStudio (version 1.1.456) using the lme4 package (version 1.1.17; Bates, Maechler, Bolker & Walker, 2015). Because it is theoretically and computationally difficult to estimate degrees of freedom in mixed-effects models, we analyzed t-scores by assuming a Gaussian distribution (Mirman, 2014). Therefore, t-values $> \pm 1.96$ were considered significant.

3. Results

Children's looking patterns are presented in Figures 1 & 2 and GCA results are included in Table 2. These results are model predictions of changes in word recognition accuracy over time *without* controlling for individual differences in children's PLS-5 auditory comprehension scores. These analyses test whether children with ASD, as a group, use coarticulation to incrementally process speech. They do not, however, control for differences in receptive language ability, which was lower for children in the ASD group compared to children in the TD group.

Children recognized the familiar words in both conditions. Collapsing across groups and conditions, children's accuracy in fixating the target image was significantly greater than chance [intercept b=0.21, t=5.65, p<.001] and increased throughout the window [linear b=0.75, t=7.13, p<.001]. The increase in children's accuracy was delayed at the onset of the window [cubic b=0.11, t=5.43, p<.001] and accelerated towards the end of the window [quadratic b=0.32, t=8.69, p<.001].

Children's word recognition accuracy, however, differed noticeably between conditions [quadratic time x condition b=0.18, t=2.35, p<.05; cubic time x condition b=0.11, t=2.54, p<.05]. Children's fixations deviated from chance significantly later on trials without coarticulation [Neutral cubic b=.16] than trials with coarticulation [Facilitating cubic b=.05]. Moreover, the increase in children's fixations to the target object accelerated later on trials without coarticulation [Neutral quadratic b=0.41] than trials with coarticulation [Facilitating quadratic b=0.23].

Crucially, the effect of condition on word recognition accuracy did not differ between groups [p's > 0.39]. Taken together, these results indicate that coarticulation facilitated word recognition and that this facilitation was the same for children in the ASD and TD groups.

In order to control for group differences in receptive language ability, we refit our GCA including children's raw auditory comprehension scores on the PLS-5 (mean-centered) as a covariate. These analyses test whether the effect of coarticulation on speech processing varies between children in the ASD and TD groups who are matched in receptive language ability. These results, however, only apply to children with average receptive language ability (raw score of 29.1 on the PLS-5 AC subtest) and so cannot be generalized to all children with ASD (or TD). Children's looking patterns are presented in Figure 2 and GCA results are included in Table 3.

Collapsing across groups and conditions, a child with average receptive language abilities was significantly greater than chance in their accuracy in fixating the target image [intercept b=0.20, t=5.86, p<.001]. Their accuracy increased significantly from the beginning to the end of the window [linear b=0.74, t=9.49, p<.001]. The increase in accuracy was delayed at the onset of the window [cubic b=0.11, t=5.27, p<.001] and accelerated towards the end of the window [quadratic b=0.31, t=8.85, p<.001].

For a child with average receptive language ability, word recognition accuracy differed between conditions [quadratic time x condition b=0.17, t=2.29, p<.05; cubic time x condition b=0.09, t=2.15, p<.05]. Their fixations deviated from chance significantly later on trials without coarticulation [Neutral cubic b=0.15] than trials with coarticulation [Facilitating cubic b=0.06]. Moreover, the increase in their fixations to the target object accelerated later on trials without coarticulation [Neutral quadratic b=0.40] than trials with coarticulation [Facilitating quadratic b=0.23].

Crucially, the effect of condition on word recognition accuracy for children with average receptive language ability differed between groups [quadratic time x condition x group b= −0.37, t=−2.35, p<.05]. The effect of coarticulation on word recognition accuracy was significantly larger for a child with average receptive language ability in the ASD group [quadratic time x condition b=0.35] than for a child with average receptive language ability in the TD group [quadratic time x condition b=−0.02]. To further examine this interaction, we re-centered our model on each condition. Word recognition accuracy did not differ between groups on trials without coarticulation [Neutral p's > 0.15] but did differ on trials with coarticulation [Facilitating quadratic:Group b=0.21, t=1.99, p<.05]. Fixations to the target object accelerated later on trials with coarticulation for children with average receptive language ability in the TD group [quadratic b=0.33] compared to the ASD group [quadratic b=0.13]. For trials with coarticulation, there were also differences between groups in their overall word recognition accuracy [Facilitating intercept:Group b=−0.21, t=−2.07, p<.05], such that children with average receptive language ability were overall more accurate in the ASD group [b=0.35] than the TD group [b=0.14].

Finally, the extent to which coarticulation affects word recognition accuracy was moderated by individual differences in children's receptive language ability [quadratic time x condition

x PLS b=0.03, t=3.96, p<.001; cubic time x condition x PLS b=.01, t=2.19, p<.05]. To explore this interaction, we refit the model to examine the effect of coarticulation for children with receptive language abilities 1 SD below the mean (raw PLS score l~ 18) and 1 SD above the mean (raw PLS score ~ 40). Recall that for children with average receptive language abilities (raw PLS ~ 29), fixations to the target image deviated from chance later [cubic time x condition b=0.09] and accelerated later [quadratic time x condition b=0.17] for trials without coarticulation compared to trials with coarticulation. For children with higher receptive language abilities, the increase from chance is even more delayed [cubic time x condition b=0.20, t=3.19, p<.01] and the later acceleration steeper [quadratic time x condition b=0.48, t=4.6, p<.001]. For children with lower receptive language abilities, however, there was not a significant effect of coarticulation on word recognition accuracy [quadratic time x condition b=−0.13, t=−1.2, p=0.22; cubic time x condition b=−0.01, t= −0.01, p=.93].

When controlling for group differences in receptive language ability, we again observed that coarticulation facilitates word recognition. This facilitation, however, was significantly stronger for children with ASD compared to children with TD. This discrepancy between our models that did versus did not control for receptive language ability is important in understanding variability in speech processing between children with ASD and will be examined further in the Discussion.

4. Discussion

We used a looking-while-listening (LWL) experiment to compare how coarticulation affects speech processing for children with ASD and children who are TD. On each trial, children saw images of two familiar objects displayed on a screen. They then heard a sentence labelling one of the objects (e.g., *Find the ball*). On Facilitating trials, the determiner *the* contained coarticulatory information about the onset of the target noun (e.g., *the*^{b}). On Neutral trials, the determiner *the* did not contain any coarticulatory information (e.g., *the*^{ə}). We found that coarticulation facilitated speech processing – children were faster to look at the target object on Facilitating compared to Neutral trials. When the groups were *not* matched in receptive language ability, we found that word recognition for children in both groups was equally facilitated by coarticulation. When controlling for group differences in receptive language ability, however, we found that the effect of coarticulation on speech processing was stronger for children with ASD compared to children who are TD. Finally, the effect of coarticulation on speech processing was stronger for children with better receptive language ability.

Taken together, our findings compellingly demonstrate that children with ASD use phonological information to incrementally process speech. These results are consistent with prior research suggesting that children with ASD use semantic information to incrementally process speech (Bavin et al., 2016; Brock et al., 2008; Hahn et al., 2015; Venker et al. 2019). The current results extend this literature to demonstrate that, like their TD peers, children with ASD may be able to use many different types of information (not just semantic cues) to support incremental language processing.

Additionally, our findings suggest that children with ASD and children who are TD may vary in their ability to use different types of information to incrementally process speech. Prior research has shown that children with ASD and children who are TD are *equally* able to use semantic information to incrementally process speech when both groups are matched in language ability – either by including only high-functioning children with ASD (Bavin et al., 2016; Hahn et al., 2015) or by including children with below average language ability in the TD group (Brock et al., 2008). We found, however, that children with ASD were *better* at using phonological information to incrementally process speech than children who are TD when controlling for differences in language ability. This is consistent with the advantage children with ASD demonstrate in other phonological tasks (Henderson et al, 2014; Nadig & Mulligan, 2017; Norbury et al., 2010) and more broadly with theories of autism (Mottron & Burack, 2001).

The observed differences between our groups in incremental speech processing, however, must be interpreted with caution. Children with ASD can either be matched to children with TD in chronological age or in language ability. We chose the latter – mismatching our groups in chronological age to match them in language ability. It is therefore possible that children in the ASD group may have been more sensitive to coarticulation, not because of anything related to their diagnosis, but rather because they are older. Previous research has demonstrated that age-related improvements in spoken word recognition continue well into adolescence for children who are TD (Rigler, Farris-Trimble, Geriner, Walker, Tomblin, & McMurray, 2015). In exploratory analyses (see Supplementary materials) we found that age and receptive language ability (PLS) were correlated for children in the TD Group and both factors were associated with incremental processing – older children and children with higher PLS scores were more affected by coarticulation. For children in the ASD Group, however, age and receptive language ability were *not* correlated and only the PLS was associated with incremental processing. These results suggest that differences in chronological age between the ASD and TD groups do not account for the differences in incremental processing. Moreover, improvements in language ability may account for age-related improvements in spoken word recognition observed for children with TD. The current experiment, however, was not designed to explicitly test these hypotheses and this remains an important topic for future research.

Finally, our results reveal that there is significant heterogeneity in the extent to which children use coarticulation to incrementally process speech and that this variability is associated with individual differences in receptive language ability. A major limitation of research on language development in children with ASD is that most of the findings are limited to high-functioning children, either because of task demands or decisions to match groups on verbal ability. Children with ASD, however, vary extensively in their language abilities (Eigsti, de Marchena, Schuh, & Kelly, 2011; Georiades et al., 2013; Kjelgaard & Tager-Flusberg, 2001; Pickles et al., 2014; Tager-Flusberg & Kasari, 2013; Wiggins et al., 2017). A strength of the current research is that our methods allowed us to include children with ASD with a wide range of verbal abilities. Although we found evidence at the group level that children used coarticulation to incrementally process speech, this does *not* mean that all children benefitted from coarticulatory cues. Indeed, we found that coarticulation did *not* affect speech processing for children with below average receptive language ability. This

association between receptive language ability and incremental processing also explains why the effect of coarticulation did *not* vary between groups without controlling for receptive language ability. Children in the ASD group had on average lower receptive language ability than children in the TD group.

The positive association between children's language ability and the effect of coarticulation is consistent with prior work examining incremental speech processing for both typically developing children between 2 and 10 years of age (Borovsky et al., 2012; Borovsky & Creel, 2014; Lew-Williams & Fernald, 2007; Mani & Huettig, 2012) and a group of children with ASD between 4 and 5 years of age with significant heterogeneity in language ability (Venker et al., 2019). Thus, increases in language ability are associated with improvements in incremental processing across a wide variety of ages and language abilities. It remains unclear, however, whether this association is causal and in which direction. It may be that improving incremental processing leads children to have larger vocabularies (by making them more successful at learning new words) or that increasing the size of children's vocabulary improves their ability to incrementally process speech (by improving speech processing speed more generally).

Superficially, our results and others demonstrating that children with ASD incrementally process speech may seem to contradict theories which propose that children with ASD have compromised prediction skills (e.g., Pellicano & Burr, 2012; Sinha et al., 2014; Van de Cruys et al., 2014). Our results, however, do not necessarily contradict these theories. Sinha and colleagues (2014) hypothesize that children with ASD may be inaccurate in estimating conditional probabilities over different time scales. Van de Cruys and colleagues (2014) propose that children with ASD have intact abilities to generate predictions and assess errors, but are inflexible in their response to prediction errors (i.e., failing to ignore prediction errors in noisy and unpredictable environments). Thus, children with ASD may do well in situations that are deterministic and exact (Mottron et al., 2013), but not in situations that are probabilistic and inexact. Coarticulation involves regularities at very short intervals (on the millisecond scale) in situations that are more deterministic and less probabilistic. It may be the case that children with ASD are able to generate predictions when processing speech better than children with TD, while using predictive information less successfully in other contexts and at other time-scales.

Our findings are part of an emerging field of research examining how children with ASD process speech. Past research has found that children with ASD use semantic information to incrementally process speech (Bavin et al., 2016; Brock et al., 2008; Hahn et al., 2015; Venker et al. 2019). In the current study, children with ASD also use phonological information – specifically coarticulation – to incrementally process speech. In fact, the ability to exploit rapid speech cues during lexical processing may be an area of relative strength for children with ASD as compared to their TD peers.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements:

References

American Psychiatric Association. (2013). Diagnostic and Statistical Manual of Mental Disorders (5th ed.). Washington, D.C.: American Psychiatric Association.

Arunachalam S, & Luyster RJ (2016). The integrity of lexical acquisition mechanisms in autism spectrum disorders: A research review. Autism Research, 9, 810–828. [PubMed: 26688218]

Barr DJ (2008). Analyzing 'visual world' eyetracking data using multilevel logistic regression. Journal of Memory and Language, 59(4), 457–474.

Bavin EL, Kidd E, Prendergast L, Baker E, Dissanayake C, & Prior M. (2014). Severity of autism is related to children's language processing. Autism Research, 7, 687–694. [PubMed: 25262588]

Borovsky A, & Creel SC (2014). Children and adults integrate talker and verb information in online processing. Developmental Psychology, 50(5), 1600. [PubMed: 24611671]

Borovsky A, Elman JL, & Fernald A. (2012). Knowing a lot for one's age: Vocabulary skill and not age is associated with anticipatory incremental sentence interpretation in children and adults. Journal of Experimental Child Psychology, 112(4), 417–436. [PubMed: 22632758]

Brock J, Norbury C, Einav S, & Nation K. (2008). Do individuals with autism process words in context? Evidence from language-mediated eye-movements. Cognition, 108, 896–904. [PubMed: 18692181]

Chang F, Dell GS, & Bock K. (2006). Becoming syntactic. Psychological Review, 113(2), 234–272. [PubMed: 16637761]

Christiansen MH, & Chater N. (2016). The now-or-never bottleneck: A fundamental constraint on language. Behavioral and Brain Sciences. Advance online publication.

Cho S-J, Brown-Schmidt S, Lee W-L (2018) Autoregressive generalized linear mixed effect models with crossed random effects: An application to intensive binary time series eye-tracking data. Psyhometrika, 83, 751–771.

Daniloff R, & Hammarberg ROBERT (1973). On defining coarticulation. Journal of Phonetics, 1(3), 239–248.

Dell GS, & Chang F. (2013). The P-chain: Relating sentence production and its disorders to comprehension and acquisition. Philosophical Transactions of the Royal Society of London. Series B, Biological Science. Advance online publication.

Bates D, Maechler M, Bolker B, & Walker S. (2015). Fitting linear mixed-effects models using lme4. Journal of Statistical Software, 67(1), 1–48.

Eigsti IM, de Marchena AB, Schuh JM, & Kelley E. (2011). Language acquisition in autism spectrum disorders: A developmental review. Research in Autism Spectrum Disorders, 5(2), 681–691.

Elman JL (1990). Finding structure in time. Cognitive Science, 14, 179–211.

Fernald A, Zangl R, Portillo AL, & Marchman VA (2008). Looking while listening: Using eye movements to monitor spoken language. In Sekerina IA, Fernández EM, and Clahsen H. (Eds.), Developmental psycholinguistics: On-line methods in children's language processing (pp. 97–135). Amsterdam: John Benjamins Publishing.

Georgiades S, Szatmari P, Boyl M, Hanna S, Duku E, Zwaigenbaum L, Bryson S, Fombonne E, Volden J, Mirenda P, Smith I, Roberts W, Vaillancourt T, Waddell C, Bennett T, Thompson A, & Pathways in ASD Study Team. Investigating phenotypic heterogeneity in children with autism spectrum disorder: A factor mixture modeling approach. The Journal of Child Psychology and Psychiatry, 54, 206–215.

Gow DW, & McMurray B. (2007). Word recognition and phonology: The case of English coronal place assimilation. Papers in Laboratory Phonology, 9(173–200).

Hahn N, Snedeker J, & Rabagliati H. (2015). Rapid linguistic ambiguity resolution in young children with autism spectrum disorder: Eye tracking evidence for the limits of weak central coherence. Autism Research, 8, 717–726. [PubMed: 25820816]

Happ F, & Booth R. (2008). The power of the positive: Revisiting weak coherence in autism spectrum disorders. Quarterly Journal of Experimental Psychology, 61, 50–63.

Happ F, & Frith U. (2006). The weak coherence account: Detail- focused cognitive style in autism spectrum disorders. Journal of Autism and Developmental Disorders, 36, 5–25. [PubMed: 16450045]

Henderson L, Powell A, Gaskell MG, & Norbury C. (2014). Learning and consolidation of new spoken words in autism spectrum disorder. Developmental Science, 17, 858–871. [PubMed: 24636285]

Huang Y. & Snedeker J. (2020). Evidence from the visual world paradigm raises questions about unaccusativity and growth curve analyses. Cognition, 200, 104251. [PubMed: 32416399]

Kidd C, White KS, & Aslin RN (2011). Toddlers use speech disfluencies to predict speakers' referential intentions. Developmental Science, 14(4), 925–934. [PubMed: 21676111]

Kjelgaard MM, & Tager-Flusberg H. (2001). An investigation of language impairment in autism: Implications for genetic sub- groups. Language and Cognitive Processes, 16, 287–308. [PubMed: 16703115]

Lew-Williams C, & Fernald A. (2007). Young children learning Spanish make rapid use of grammatical gender in spoken word recognition. Psychological Science, 18(3), 193–198. [PubMed: 17444909]

Lord C, Rutter M, DiLavore PC, Risi S, Gotham K, & Bishop S. (2012). Autism Diagnostic Observation Schedule, Second Edition (ADOS-2) Manual (Part 1): Modules 1–4. Torrence: Western Psychological Services.

Lukyanenko C, & Fisher C. (2016). Where are the cookies? Two-and three-year-olds use number-marked verbs to anticipate upcoming nouns. Cognition, 146, 349–370. [PubMed: 26513355]

Mahr T, McMillan BT, Saffran JR, Ellis Weismer S, & Edwards J. (2015). Anticipatory coarticulation facilitates word recognition in toddlers. Cognition, 142, 345–350. [PubMed: 26072992]

Mani N, & Huettig F. (2012). Prediction during language processing is a piece of cake—But only for skilled producers. Journal of Experimental Psychology: Human Perception and Performance, 38(4), 843. [PubMed: 22774799]

Mattys SL, White L, & Melhorn JF (2005). Integration of multiple speech segmentation cues: a hierarchical framework. Journal of Experimental Psychology: General, 134(4), 477. [PubMed: 16316287]

McClelland JL, & Elman JL (1986). The TRACE model of speech perception. Cognitive Psychology, 18(1), 1–86. [PubMed: 3753912]

Mirman D. (2014). Growth curve analysis and visualization using R. Boca Raton, FL: CRC Press.

Mottron L, Bouvet L, Bonnel A, Samson F, Burack JA, Dawson M, & Heaton P. (2013). Veridical mapping in the development of exceptional autistic abilities. Neuroscience & Biobehavioral Reviews, 37, 209–228. doi:10.1016/j.neubiorev.2012.11.016 [PubMed: 23219745]

Mottron L, & Burack J. (2001). Enhanced perceptual functioning in the development of autism. In Burack J, Charman T, Yirmiya N, & Zelazo PR (Eds.), The Development of Autism: Perspectives from Theory and Research (pp. 131–148). Mahwah, NJ: Lawrence Erlbaum Associates.

Mullen EM. (Ed.). (1995). Mullen Scales of Early Learning Manual. Minneapolis, MN: AGE Edition.

Nadig A, & Mulligan A. (2017). Intact non-word repetition and similar error patterns in language-matched children with autism spectrum disorders: A pilot study. Journal of Communication Disorders, 66, 13–21. [PubMed: 28349875]

Norbury CF, Griffiths H, & Nation K. (2010). Sound before meaning: Word learning in autistic disorders. Neuropsycologia, 48, 4012–4019.

Pellicano E, & Burr D. (2012). When the world becomes ―too real‖: A Bayesian explanation of autistic perception. Trends in Cognitive Sciences, 16, 504–510. [PubMed: 22959875]

Pickering MJ, & Garrod S. (2013). An integrated theory of language production and comprehension. Behavioral and Brain Sciences, 36, 329–347.

Pickles A, Anderson D, & Lord C. (2014). Heterogeneity and plasticity in the development of language: A 17-year follow-up of children referred early for possible autism. Journal of Child Psychology and Psychiatry, 55, 1354–1362. [PubMed: 24889883]

Pomper R, Ellis Weismer S, Saffran J, & Edwards J. (2019). Specificity of phonological representations for children with autism spectrum disorder. Journal of Autism and Developmental Disorders, 49(8), 3351–3363. 10.1007/s10803-019-04054-5. [PubMed: 31098924]

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Reuter T, Borovsky A, & Lew-Williams C. (2019). Predict and redirect: Prediction errors support children's word learning. Developmental psychology.

Rigler H, Farris-Trimble A, Greiner L, Walker J, Tomblin JB, & McMurray B. (2015). The slow developmental time course of real-time spoken word recognition. Developmental Psychology, 51, 1690–1703. [PubMed: 26479544]

Robins DL, Fein D, Barton ML, & Green JA (2001). The modified checklist for autism in toddlers: An initial study investigating the early detection of autism and pervasive developmental disorders. Journal of Autism and Developmental Disorders, 31, 131–144. [PubMed: 11450812]

Rutter M, LeCouteur A, & Lord C. (2003). Autism Diagnostic Interview-Revised. Los Angeles: Western Psychological Service.

Salverda AP, Kleinschmidt D, & Tanenhaus MK (2014). Immediate effects of anticipatory coarticulation in spoken-word recognition. Journal of Memory and Language, 71(1), 145–163. [PubMed: 24511179]

Sinha P, Kjelgaard MM, Gandhi TK, Tsourides K, Cardinaux AL, Pantazis D, et al. (2014). Autism as a disorder of pre- diction. Proceedings of the National Academy of Sciences, 111, 15220–15225.

Tager-Flusberg H, & Kasari C. (2013). Minimally verbal school-aged children with autism spectrum disorder: The neglected end of the spectrum. Autism Research, 6, 468–478. [PubMed: 24124067]

Van de Cruys S, Evers K, Van der Hallen R, Van Eylen L, Boets B, De-Wit L, & Wagemans J. (2014). Precise minds in uncertain worlds: Predictive coding in autism. Psychological Review, 121, 649–675. [PubMed: 25347312]

Venker CE, Edwards J, Saffran JR, & Ellis Weismer S. (2019). Thinking ahead: Incremental language processing is associated with receptive language abilities in preschoolers with autism spectrum disorder. Journal of Autism and Developmental Disorders, 49(3), 1011–1023. [PubMed: 30390172]

Venker CE, Haebig E, Edwards J, Saffran JR, & Ellis Weismer S. (2016). Brief report: Early lexical comprehension in young children with ASD: Comparing eye-gaze Methodology and parent report. Journal of Autism and Developmental Disorders, 46, 2260–2266. [PubMed: 26883646]

Venker CE, Pomper R, Mahr T, Edwards J. Saffran J, & Ellis Weismer S (2019). Comparing automatic eye tracking and manual gaze coding methods in young children with autism spectrum disorder. Autism Research, 1–13.

Wiggins LD, Tian LH, Levy SE, Rice C, Lee L-C. Schieve L, Pandey J, Daniels J, Blaskey L, Hepburn S, Landa R, Edmondson-Pretzel R, & Thompson W. (2017). Homogenous subgroups of young children with autism improve phenotypic characterization in the study to explore early development. Journal of Autism and Developmental Disorders, 47, 3634–3645. [PubMed: 28879490]

Zimmerman IL, Steiner VG, & Pond RE (2011). Preschool language scales, Fifth Edition. San Antonio: The Psychological Corporation.
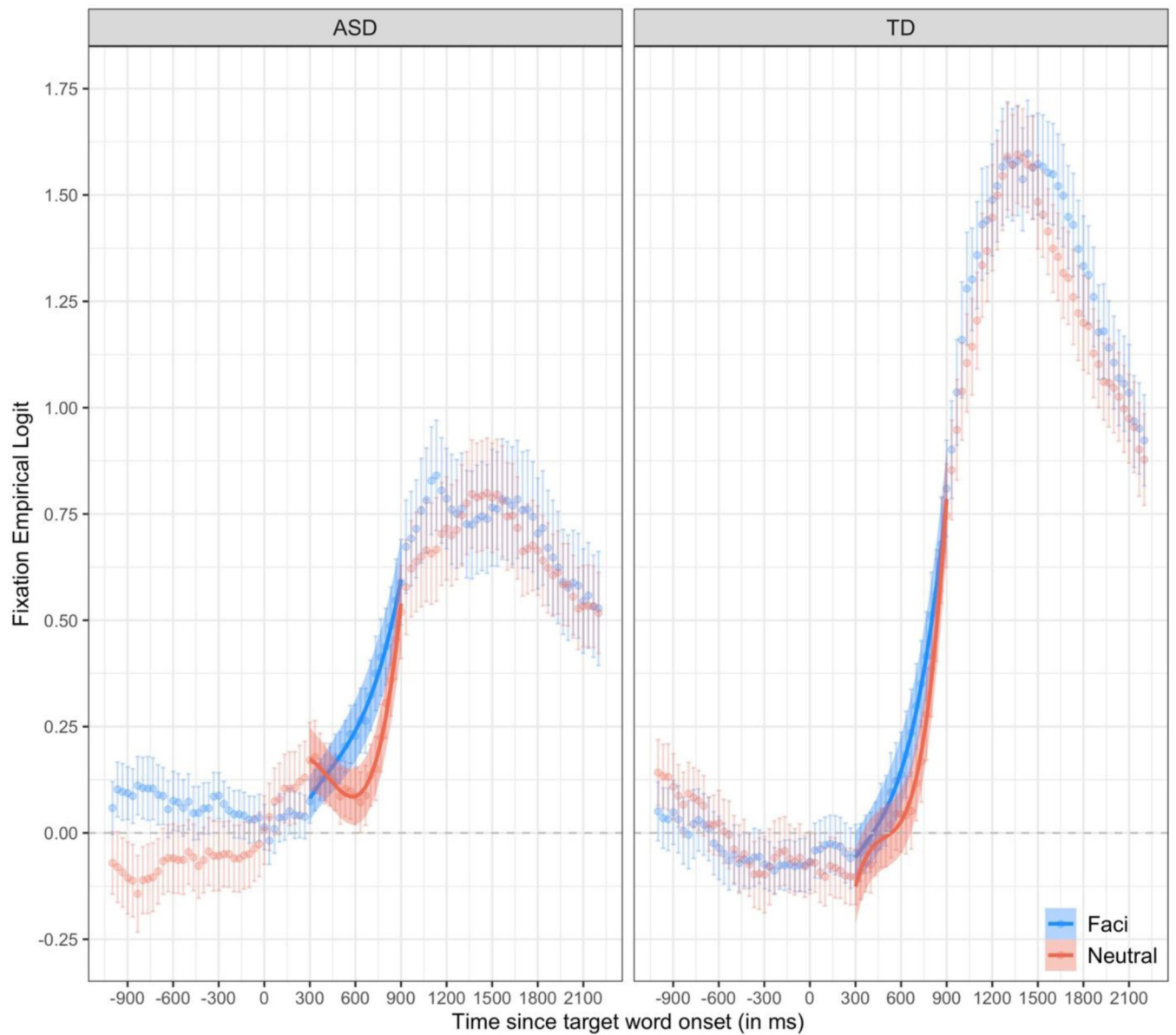
Figure 1. Time course of children's word recognition accuracy.
The empirical log-odds of fixating the target image over time are plotted for trials with
coarticulation (Facilitating in blue) and trials without coarticulation (Neutral in red). Data
points are observed behavioral data averaged across children. Lines are growth curve model
fits. Ribbons around the lines represent ± 1 SE. The dashed horizontal line at 0 represents
chance (i.e., equal likelihood of fixating the target and distractor images).

Figure 2. Time course of children's word recognition accuracy during the critical window.
The empirical log-odds of fixating the target image over time are plotted for trials with
coarticulation (Facilitating in blue) and trials without coarticulation (Neutral in red). Data
points are observed behavioral data averaged across children. Lines are growth curve model
fits. Ribbons around the lines represent ± 1 SE. The dashed horizontal line at 0 represents
chance (i.e., equal likelihood of fixating the target and distractor images).
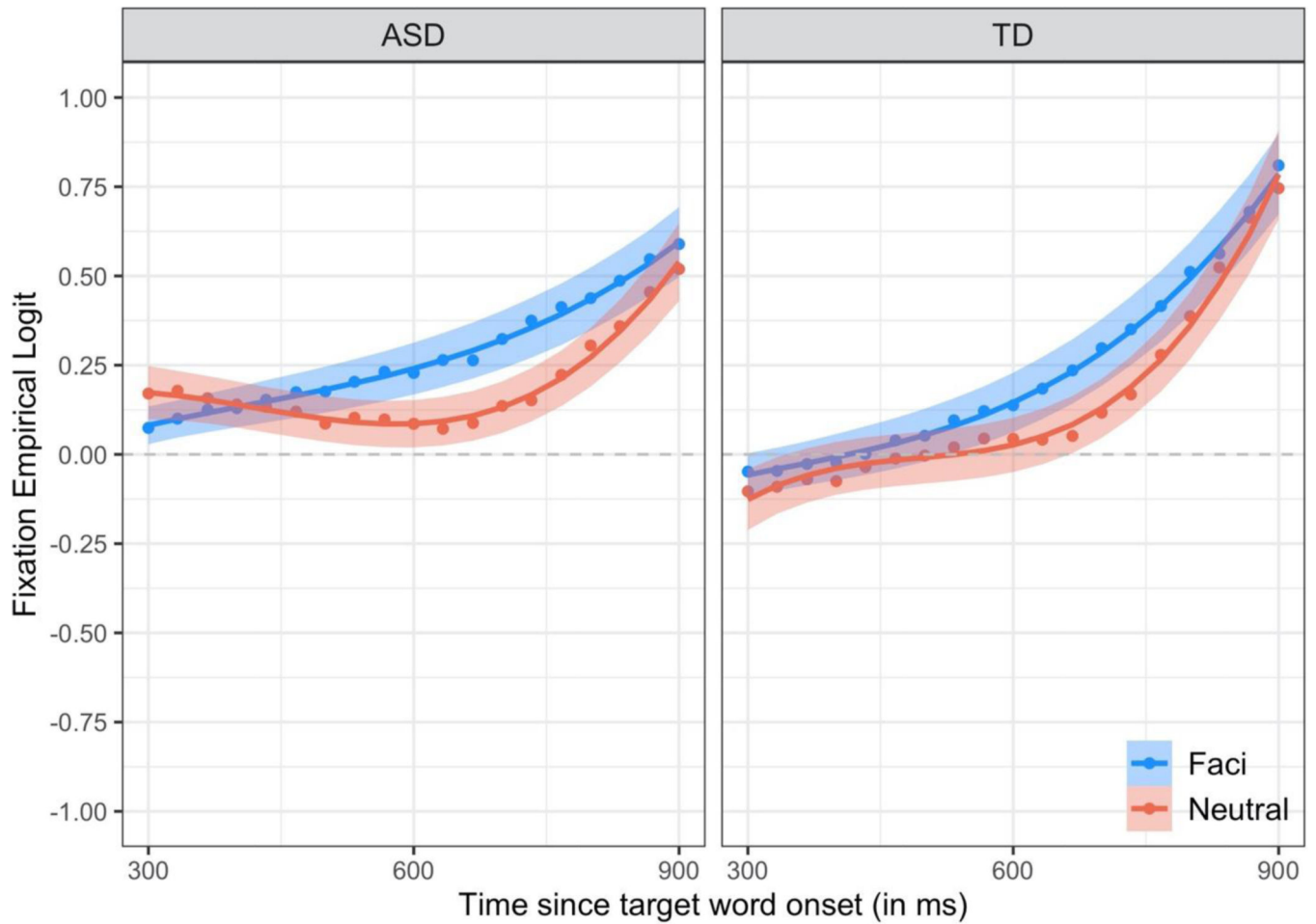
Figure 3. Time course of children's word recognition accuracy during the critical window controlling for receptive language ability.

The empirical log-odds of fixating the target image across time are plotted for trials with coarticulation (Facilitating trials in blue) and trials without coarticulation (Neutral trials in red). The lines are growth curve model fits for a child with an average score on the Auditory Comprehension scale of the PLS-5. Ribbons around the lines represent ± 1 SE. The dashed horizontal line at 0 represents chance (i.e., equal likelihood of fixating the target and distractor images).
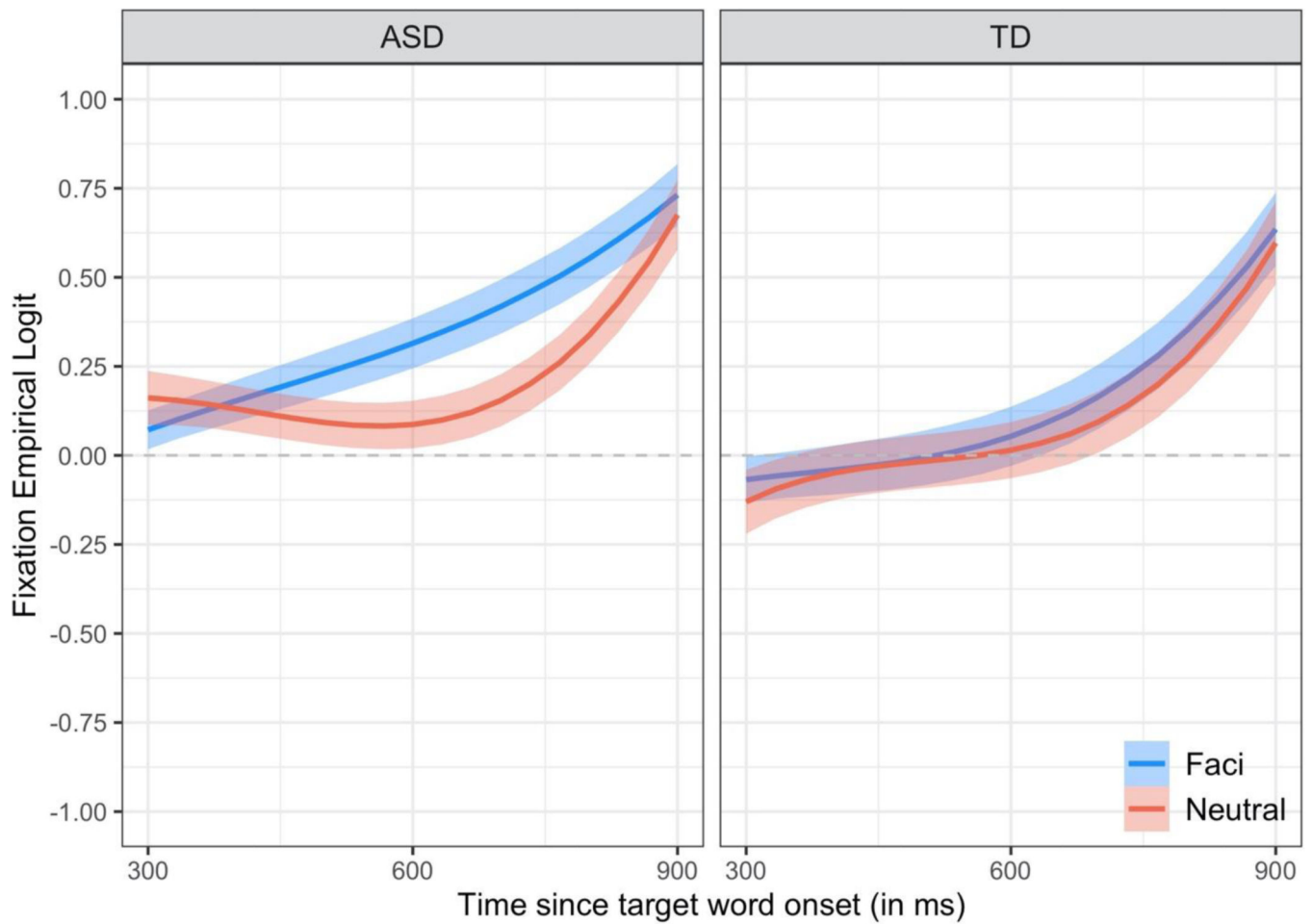
Table 1.

Participant characteristics for the autism spectrum disorders (ASD) and typically developing (TD) groups.

Receptive Language refers to children's score on the Auditory Comprehension scale of the Preschool Language Scales, 5th edition. Nonverbal Cognition refers to children's score on the Visual Reception scale of the Mullen Scales of Early Learning. Autism severity refers to children's standardized calibrated severity scores on the Autism Diagnostic Observation Schedule, 2nd edition.

	ASD (*n* = 58)		TD (*n* = 44)		
	Mean (SD)	Range	Mean (SD)	Range	
Age (months)	43.79 (3.52)	37–50	26.66 (5.82)	18–36	*
Maternal education (years)	13.78 (2.11)	10–22	17.34 (2.27)	12–24	*
Autism severity	8.19 (1.63)	4–10	-	-	
Receptive Language					
PLS AC Raw	25.63 (11.89)	10–56	33.71 (7.96)	19–49	*
PLS AC Standard Score	66.21 (20.64)	50–133	111.38 (14.14)	81–137	*
Nonverbal Cognition					
Mullen VR Raw	29.84 (8.49)	14–48	32.09 (7.11)	20–49	
Mullen VR T-Score	29.6 (14.0)	19–63	59.9 (11.5)	39–81	*
Race/Ethnicity					
White	50		43		
Black	2		0		
Asian	0		0		
Native American/Alaska Native	1		0		
More than one race	5		1		
Hispanic or Latino	4		0		

*
Indicates group differences at $p < 0.05$.

Table 2.

Growth curve model estimates and tests of significance.

Results for model fit using the full sample of children without controlling for group differences in receptive language ability.

Fixed Effects	Estimate	SE	t	p	
(Intercept)	0.21	0.04	5.65	<0.01	*
ot1	0.75	0.09	8.13	<0.01	*
ot2	0.32	0.04	8.69	<0.01	*
ot3	0.11	0.02	5.43	<0.01	*
Condition	-0.09	0.06	-1.38	0.17	
ot1:Condition	-0.18	0.12	-1.57	0.12	
ot2:Condition	0.18	0.08	2.35	0.02	*
ot3:Condition	0.11	0.04	2.54	0.01	*
Group	-0.05	0.07	-0.67	0.5	
ot1:Group	0.51	0.18	2.77	0.01	*
ot2:Group	0.11	0.07	1.49	0.14	
ot3:Group	0.06	0.04	1.53	0.13	
Condition:Group	0.00	0.13	0.02	0.98	
ot1:Condition:Group	0.20	0.24	0.83	0.4	
ot2:Condition:Group	-0.13	0.16	-0.82	0.41	
ot3:Condition:Group	0.08	0.09	0.86	0.39	

Table 3.

Growth curve model estimates and tests of significance.Fixed effects abbreviations for the orthogonal time terms are linear (ot1), quadratic (ot2), and cubic (ot3). PLS is children's mean-centered, raw score on the Auditory Comprehension scale of the PLS-5.

Fixed Effects	Estimate	SE	t	p	
(Intercept)	0.20	0.03	5.86	<0.01	*
ot1	0.74	0.08	9.49	<0.01	*
ot2	0.31	0.04	8.85	<0.01	*
ot3	0.11	0.02	5.27	<0.01	*
Condition	-0.09	0.07	-1.38	0.17	
ot1:Condition	-0.19	0.12	-1.57	0.12	
ot2:Condition	0.17	0.07	2.29	0.02	*
ot3:Condition	0.09	0.04	2.15	0.03	*
Group	-0.16	0.07	-2.23	0.03	*
ot1:Group	0.13	0.17	0.8	0.42	
ot2:Group	0.02	0.08	0.29	0.77	
ot3:Group	0.03	0.04	0.74	0.46	
PLS	0.02	<0.01	4.83	<0.01	*
ot1:PLS	0.05	0.01	6.23	<0.01	*
ot2:PLS	0.01	<0.01	3.43	<0.01	*
ot3:PLS	0.00	<0.01	1.96	0.05	+
Condition:Group	0.09	0.14	0.68	0.50	
ot1:Condition:Group	0.24	0.26	0.91	0.36	
ot2:Condition:Group	-0.37	0.16	-2.35	0.02	*
ot3:Condition:Group	-0.01	0.09	-0.1	0.92	
Condition:PLS	-0.01	0.01	-1.56	0.12	
ot1:Condition:PLS	-0.01	0.01	-0.61	0.54	
ot2:Condition:PLS	0.03	0.01	3.96	<0.01	*
ot3:Condition:PLS	0.01	<0.01	2.19	0.03	*

*
P values <.05 are indicated with a.