

RESEARCH ARTICLE OPEN ACCESS

Acoustic Exaggeration Enhances Speech Discrimination in Young Autistic Children

Luodi Yu^{1,2}  | Lizhi Ban^{1,2} | Aiwen Yi³ | Jing Xin⁴ | Suping Li⁵ | Suiping Wang²  | Laurent Mottron⁵

¹Center for Autism Research, School of Education, Guangzhou University, Guangzhou, China | ²Philosophy and Social Science Laboratory of Reading and Development in Children and Adolescents, Ministry of Education, South China Normal University, Guangzhou, China | ³Department of Pediatrics, Guangdong Provincial Key Laboratory of Major Obstetric Diseases, Guangdong Provincial Clinical Research Center for Obstetrics and Gynecology, Guangdong-Hong Kong-Macao Greater Bay Area Higher Education Laboratory of Maternal-Fetal Joint Medicine, The Third Affiliated Hospital of Guangzhou Medical University, Guangzhou, China | ⁴Foshan Clinical Medical School, Guangzhou University of Chinese Medicine, Foshan, China | ⁵Psychiatry and Addictology Department, CIUSSS-NIM Research Center, University of Montreal, Montreal, Quebec, Canada

Correspondence: Luodi Yu (yuluodi@gzhu.edu.cn) | Laurent Mottron (laurent.mottron@umontreal.ca)

Received: 26 July 2024 | **Revised:** 14 December 2024 | **Accepted:** 18 December 2024

Funding: This research was supported by the National Natural Science Foundation of China (31900775) and Plan on enhancing scientific research in Guangzhou Medical University (2024SRP123).

Keywords: auditory processing | autism spectrum disorders | EEG | multivariate pattern analysis | speech processing

ABSTRACT

Child-directed speech (CDS), which amplifies acoustic and social features of speech during interactions with young children, promotes typical phonetic and language development. In autism, both behavioral and brain data indicate reduced sensitivity to human speech, which predicts absent, decreased, or atypical benefits of exaggerated speech signals such as CDS. This study investigates the impact of exaggerated fundamental frequency (F0) and voice-onset time on the neural processing of speech sounds in 22 Chinese-speaking autistic children aged 2–7 years old with a history of speech delays, compared with 25 typically developing (TD) peers. Electroencephalography (EEG) data were collected during passive listening to exaggerated and non-exaggerated syllables. A time-resolved multivariate pattern analysis (MVPA) was used to evaluate the potential effects of acoustic exaggeration on syllable discrimination in terms of neural decoding accuracy. For non-exaggerated syllables, neither the autism nor the TD group achieved above-chance decoding accuracy. In contrast, for exaggerated syllables, both groups achieved above-chance decoding, indicating significant syllable discrimination, with no difference in accuracy between the autism and TD groups. However, the temporal generalization patterns in the MVPA results revealed distinct neural mechanisms supporting syllable discrimination between the groups. Although the TD group demonstrated a left-hemisphere advantage for decoding and generalization, the autism group displayed similar decoding patterns between hemispheres. These findings highlight the potential of selective acoustic exaggeration to support speech learning in autistic children, underscoring the importance of tailored, sensory-based interventions.

1 | Introduction

Autism spectrum disorder (ASD, henceforth referred to as autism or autistic) is a group of neurodevelopmental conditions characterized by challenges in social communication and interaction, sensory/perceptual differences, and restricted patterns

of behavior, interests, and activities (American Psychiatric Association 2013). Atypical speech and language development is a common feature among autistic individuals. Delayed speech onset in infancy is one of the earliest and most frequent signs of subsequent autism development (Pierce et al. 2019; Tager-Flusberg, Paul, and Lord 2005). It is estimated that as many as

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial](https://creativecommons.org/licenses/by-nc/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2024 The Author(s). *Autism Research* published by International Society for Autism Research and Wiley Periodicals LLC.

25% of autistic children are either nonverbal or have very limited verbal abilities when they start school (Norrelgen et al. 2015).

These challenges underscore the importance of investigating whether the factors influencing typical language acquisition are modified in autistic children. In typical development, the quantity and quality of auditory input play a pivotal role in shaping children's language acquisition. Child-directed speech (CDS) is a communicative style characterized by the use of exaggerated acoustic features and social affect when interacting with young children. It has been shown to be beneficial for typical children's phonetic learning and subsequent language development (Kuhl 2010). In autism, auditory perceptual atypicalities for social and non-social auditory material (O'Connor 2012; Samson et al. 2011) and delayed speech onset have been largely replicated in studies conducted in both Western (Pierce et al. 2019; Tager-Flusberg, Paul, and Lord 2005; Zwaigenbaum et al. 2005) and Eastern worlds (Li et al. 2018). Despite these findings, the impact of CDS on the neural processing of speech sounds in autism has been minimally investigated, yielding inconsistent evidence. The current study aims to use neural decoding methodology to investigate the effects of CDS-inspired acoustic exaggeration on speech processing in autistic children.

1.1 | CDS and Acoustic Exaggeration

In comparison to adult-directed speech (ADS), prototypical CDS exhibits distinctive spectral characteristics, including expanded vowel formant space (Cox et al. 2023; Cristia 2010; Kuhl et al. 1997) and a greater range and variation in pitch (F0) (Cox et al. 2023; Fernald and Simon 1984; Narayan and McDermott 2016), which have been observed across cultures and languages. For example, the CDS of mothers from the USA, Russia, and Sweden displays formant frequencies for cardinal vowels (/a/, /i/, /u/) that are more acoustically distant compared with those observed in ADS (Kuhl et al. 1997). CDS from mothers speaking a tonal language (i.e., Mandarin Chinese) not only accentuates vowel formant space but also amplifies F0 onset-offset range of lexical tones in comparison to ADS. Furthermore, CDS modifies acoustic speech cues in the temporal domain. Voice-onset-time (VOT) refers to the time interval between the release of an oral constriction (e.g., a lip constriction) and the onset of voicing, which is a defining feature of stop consonants. The CDS of English-speaking mothers exhibits longer VOTs for voiceless stop consonants (/p/, /t/, /k/) compared with ADS (Burnham et al. 2013; Englund 2005).

The exaggeration of vowel features is closely associated with infants' ability to discriminate speech sounds and their receptive language development (Hartman, Ratner, and Newman 2017; Kalashnikova and Burnham 2018; Liu, Kuhl, and Tsao 2003). Infants up to 9 months of age demonstrated a pronounced auditory preference for CDS in comparison to ADS (Cooper and Aslin 1990; Dunst, Simkus, and Hamby 2012). CDS has been shown to elicit infants' gaze following toward adults more effectively than ADS (Senju and Csibra 2008). This heightened attentional engagement is thought to facilitate the formation of associations between words spoken by adults and their referents (Zangl and Mills 2007), thereby promoting word learning (Fisher and Tokura 1995; Gangopadhyay and Kaushanskaya 2020). A meta-analysis by Spinelli, Fasolo, and Mesman (2017) indicates that the prosodic features of CDS, such

as higher pitch and greater pitch variability, are associated with better pre-linguistic skills (Gratier and Devouche 2011) and improved language outcomes (D'Odorico and Jacob 2006; Lyakso, Frolova, and Grigorev 2014).

1.2 | Behavioral and Cortical Phonetic Discrimination of Typical and Autistic Children

Speech perception is fundamental in scaffolding the typical development of vocabulary and grammatical abilities (Kuhl 2010; Werker and Hensch 2015). Challenges at this level are plausibly involved in speech onset delays in autism. In a study conducted by Chen and Peng (2021), a diminished categorical perception of stop consonant contrast (ba vs. pa) differing in VOT was observed among school-aged autistic children in comparison to their TD peers. The perception of consonants depends on the ability to distinguish subtle and transient acoustic differences, which requires a high degree of auditory temporal precision (Steinschneider et al. 2005; Tallal, Miller, and Fitch 1993). In the study of Kuhl et al. (2005), only TD children but not autistic children as a group displayed significant mismatch negativity (MMN) responses to consonant contrast (ba vs. wa) differing in initial formant transition duration. Furthermore, autistic children displayed slower auditory evoked response latencies and reduced neural oscillatory synchrony (Edgar et al. 2015; Gandal et al. 2010; Miron et al. 2021; Yu et al. 2018) compared with TD children. In addition, they presented challenges in discerning auditory temporal attributes, including a prolonged temporal binding window (Foss-Feig et al. 2010; Kwakye et al. 2011), and reduced MMN/magnetic mismatch fields responses to duration changes (Huang et al. 2018; Kasai et al. 2005; Lepisto et al. 2008). These basic auditory processing differences may impose constraints on the perception of consonants in these children.

Phonetic processing is typically lateralized, with the right hemisphere being mainly responsible for processing slowly changing, supra-segmental speech patterns such as tonal features, while the left hemisphere primarily manages rapidly changing, segmental features, such as stop consonants (Abrams et al. 2008; Hickok and Poeppel 2007). Atypical lateralization has been observed in multiple domains in autistic individuals, including auditory, phonetic, somatosensory, and linguistic functions (Deemyad 2022; Finch et al. 2017; Lai et al. 2024; Morrel et al. 2023).

In addition to differences in consonant perception, altered perception of lexical tones has been frequently reported in studies with Chinese-speaking autistic children. Diminished MMN—a neural signature of automatic perceptual discrimination—was observed in school-age autistic children for lexical tones (Wang et al. 2017; Yu et al. 2015; Zhang et al. 2019), despite enhanced pitch perception for nonspeech and musical materials (for a review, see Chen et al. 2022). In contrast to their TD counterparts, autistic children showed no differentiation between native (lexical tone) and non-native (linearly rising) pitch contours in their auditory neural responses, suggesting an atypical neural specialization for linguistic pitch information (Yu et al. 2021). Collectively, the aforementioned speech perception differences may result in inadequate processing of linguistically relevant speech signals, thereby affecting autistic children's acquisition of higher-order linguistic structures and functions (Yu and Wang 2021).

1.3 | Behavioral and Cortical Response to CDS in Autism Versus Typical Development

Exaggerated speech has been shown to enhance the neural responsiveness measured by event-related potentials (ERP), even in the absence of non-acoustic, multimodal, and emotional aspects of motherese. For example, infants aged 6 and 13 months exhibited enhanced N600-800 responses to audio-presented CDS relative to ADS, indicative of improved neural processing (Zangl and Mills 2007). Similarly, the neural responses of infants aged between 6 and months to the vowel /i/ showed a more prominent N250 component when the vowel formant space was exaggerated (Zhang et al. 2011). A recent study using functional near-infrared spectroscopy observed that CDS facilitated word learning and increased left frontoparietal activity in toddlers (Zhou et al. 2024).

Unlike TD children, who exhibit sensitivity and a preference for naturally recorded CDS, autistic infants and toddlers tend to show a lower preference for CDS compared with other sound forms, including speech-like sounds created by spectrally converting speech (Paul et al. 2007) and nonspeech sounds (environmental sounds and music; Pierce et al. 2023). This makes an atypical preference for CDS a candidate early marker of autism (Pierce et al. 2023; but see Droucker, Curtin, and Vouloumanos 2012). The exaggeration of vowel formant space, similar to the stimuli used in Zhang et al. (2011), did not result in an enhancement of the auditory P1 amplitude for autistic children to the same extent as it did for TD children (Chen et al. 2021). Autistic children at the group level preferred nonspeech sounds over CDS, and demonstrated a weaker MMN for consonant discrimination compared with TD children (Kuhl et al. 2005). The subgroup that exhibited a preference for CDS showed an MMN pattern that was comparable to that observed in the TD group. In contrast, the subgroup that preferred nonspeech sounds displayed an absence of MMN. These findings suggest that, despite individual differences, an auditory preference for CDS is associated with better phonetic abilities in some autistic children.

1.4 | The Current Study

The current study investigates the effectiveness and mechanisms of acoustic exaggeration on the neural processing of speech sounds in young Chinese-speaking autistic children, in comparison to their TD peers, with a focus on linguistic pitch (lexical tones) and stop consonants. The tonal nature of Chinese

allows for the independent and simultaneous manipulation of tone (a right hemisphere property) and consonant VOT (a left hemisphere property), thus enabling experimental comparisons within the same linguistic context, which is not feasible in non-tonal languages. We recorded the electroencephalography (EEG) in response to syllables with and without selective exaggerations of fundamental frequency (F0) of the lexical tone or VOT of the stop consonant. To determine whether and to what extent the acoustic exaggeration of the phonemic cues can help the autistic brain in differentiating between syllables, we used a multivariate pattern analysis (MVPA) approach. This method allows for the characterization of condition differences using neural patterns elicited by those conditions, in our case, syllable classes, such that permits the discovery of effects that may be overlooked by univariate methods (King and Dehaene 2014; Marsicano, Bertini, and Ronconi 2024). Moreover, MVPA is particularly advantageous in speech perception studies when behavioral measures in young autistic children are extremely difficult to obtain. Our general hypothesis is as follows: If the acoustic exaggeration of phonemic cues has positive effects, it would result in improved neural discrimination (decoding) of the exaggerated syllables in comparison to those of the non-exaggerated syllables, as reflected by increased MVPA classification accuracy or decoding clusters. In addition, we seek to examine how different acoustic properties—sustained/spectral (F0) versus transient/temporal (VOT)—influence neural processing in autistic and TD children by analyzing hemispheric differences. A left-dominant neural decoding pattern would imply that temporal cues (VOT) bear more importance in syllable discrimination, whereas a right-dominant pattern would suggest a greater reliance on spectral cues (F0).

2 | Methods

2.1 | Participants

Twenty-four autistic children, aged between 2 and 7 years were recruited from the pediatric department of a local hospital, and 26 TD children were recruited through flyers and personal contacts. Two children were excluded due to poor data quality (less than 60% usable for analysis), and one was excluded due to a subsequent diagnosis of chromosomal abnormalities. This resulted in 47 participants (22 autistic children, 25 TD children), matched on chronological age and sex (Table 1). The diagnoses in accordance with the Diagnostic and Statistical Manual of

TABLE 1 | Participant characteristics.

	Autism (<i>N</i> = 22)	TD (<i>N</i> = 25)	<i>t</i>	<i>p</i>
Sex	14 boys, 8 girls	16 boys, 9 girls	0.00	0.979
Age in month	47.8 (16.5, 27–80)	48.2 (12.6, 29–82)	−0.97	0.848
Mental age in month	29.4 (17.8, 8–59)	56.4 (17.0, 21–80)	−5.31	<0.001
SRS	87.32 (25.09, 41–135)	49.00 (13.91, 24–72)	6.36	<0.001
ABC	63.36 (29.32, 20–140)	—	—	—
CARS	34.45 (3.76, 30–42)	—	—	—

Note: SDs and ranges in parentheses.

Abbreviations: ABC, Autism Behavior Checklist; CARS, Childhood Autism Rating Scale; SRS, Social Responsive Scale.

Mental Disorders (DSM-5) criteria were provided by pediatricians with a minimum of 10 years of experience in diagnosing ASDs. All children in the autism group had presented speech onset delay. By the time of study participation, two of these children were non-verbal and occasionally spoke in single syllables, and the others spoke in words, phrases, or simple sentences. All children had normal hearing, as confirmed through auditory brainstem response measurements and parental reports. No genetic conditions, additional psychiatric disorders, neurodevelopmental disorders, or family history of language disorders were identified in any of the children. None were receiving medication at the time of the study. The Edinburgh Handedness Inventory was administered to 43 children. Three children with autism and two typically developing children were left-handed, while one typically developing child exhibited no clear hand dominance.

Before the EEG experiments, the receptive language abilities of the participants were evaluated using the Peabody Picture Vocabulary Test-Revised (PPVT-R) or the short form of the Chinese Communicative Development Inventory (CDI). The assessment scores were converted to mental age using CDI and PPVT-R normative data for Chinese children (Gong and Guo 1984; Sang and Liao 1990; Tardif and Fletcher 2008). The Social Responsiveness Scale (Cen et al. 2017; Constantino and Gruber 2009) was completed by caregivers for autism screening in the TD group. Five TD children scored above the cutoff (≥ 60 , 66–72), but autism was excluded after further evaluations by qualified pediatricians. The symptom load of autistic children was assessed using the Autism Behavior Checklist (ABC) and the Childhood Autism Rating Scale (CARS), with higher scores indicating greater symptoms. The ABC was completed by caregivers, and the CARS by pediatricians. Informed consent was obtained from the caregiver of each child, following a protocol approved by the local Research Ethics Committee.

2.2 | Stimuli

The vowel-only syllable /i2/ (/i/ with a rising Tone 2) and the CV syllable /p^hi1/ (with a high-level Tone 1) were used as target syllables. To create acoustically exaggerated versions of these syllables, the fundamental frequency (F_0) range of /i2/ was increased, and the VOT of /p^hi1/ was lengthened (Figure 1). These acoustic adjustments were based on lab-recorded CDS and literature on Chinese mothers' CDS (Fish et al. 2017; Tang et al. 2017;

see Supporting Information for detailed procedures). The target syllables were synthesized using a commercial text-to-speech software. Adjustments were made to create their exaggerated versions based on the determined acoustic parameters (Table 2 and Figure 1). To maintain the total length of the exaggerated syllable unchanged, the vowel length was adjusted accordingly when elongating the VOT. All sounds were standardized to a sampling rate of 44,100 Hz, a duration of 350 ms, and an intensity of 70 dB SPL.

2.3 | Procedure

Sequences of exaggerated and non-exaggerated syllables were presented via insert earphones. Each syllable was presented in four short blocks, comprising two exaggerated and two non-exaggerated blocks, in a random order. Each short block contained 50 trials, resulting in 400 trials. The inter-stimulus interval jittered between 1000 and 1200 ms. During the EEG recording, participants were seated in a comfortable position and instructed to watch a muted cartoon of their choosing, while ignoring any sounds. The EEG recording lasted approximately 12 min, with the entire EEG experiment, including preparation, taking approximately 20 min.

2.4 | EEG Recording and Analysis

The EEG was recorded using a 64-channel EGI HydroCel Geodesic System (HGSN) system and a NetAmps 300 amplifier (Electrical Geodesic Inc., Eugene, OR), at a sampling rate of 1000 Hz. The electrode impedance was kept below 50 k Ω . The Cz reference was used.

TABLE 2 | Acoustic parameters of interest of the exaggerated and non-exaggerated syllables.

	/i2/		/p ^h i1/
	F0 average (Hz)	F0 range (Hz)	VOT (ms)
Non-exaggerated	240	82	82
Exaggerated	240	131	97.6

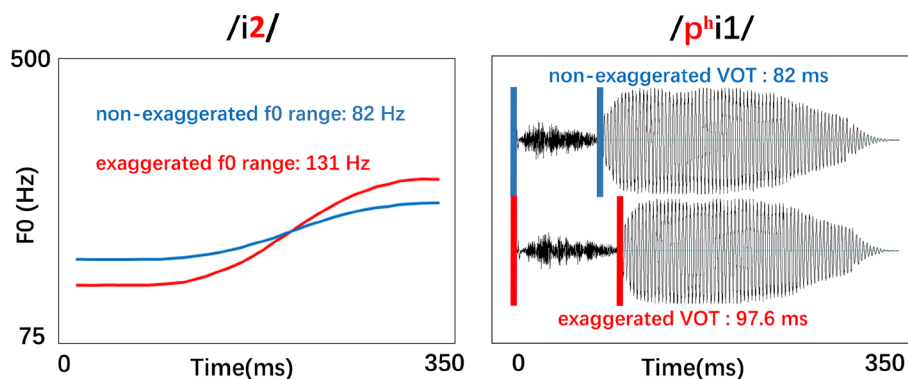


FIGURE 1 | Acoustic features of the exaggerated and non-exaggerated syllables. (Left) F0 (tonal) exaggeration; (right) VOT exaggeration.

2.4.1 | EEG Preprocessing

The EEG data preprocessing was conducted with EEGLAB toolbox (Delorme and Makeig 2004). The raw data underwent band-pass filtering at 0.5–40 Hz using a Hamming windowed sinc finite impulse response filter. Motion artifacts were removed using independent component analysis. The signals were then re-referenced to linked mastoids. Epochs of 1000 ms, including a 200-ms pre-stimulus baseline, were extracted for subsequent analysis. Trials with instantaneous amplitude exceeding $\pm 100 \mu V$ were rejected. Those participants retaining fewer than 60% of the trials were excluded from further analysis. The average trial retention rates for the four stimulus conditions in the autism group ranged from 83.6% to 85.2% (61–98 trials). In the TD group, the trial retention rates were 84.8%–87.5% (61–98 trials).

2.4.2 | Univariate Analysis: ERP Waveform Analysis

Before MVPA, a conventional waveform analysis was performed using the global field power (GFP). GFP, calculated as the SD of the ERP amplitude across all electrodes, is a measure of neural activity independent of scalp location (Lehmann and Skrandies 1984). Non-parametric cluster-based permutation tests (CBPT; Maris and Oostenveld 2007) were used to examine the effects of acoustic exaggeration on response amplitude across the entire post-stimulus epoch of the GFP waveforms.

2.4.3 | MVPA: Classification (Diagonal Decoding)

The use of MVPA on EEG/ERP data represents a powerful methodology for investigating auditory and speech processing, as it provides high temporal resolution for tracking neural dynamics, while requiring minimal a priori assumptions about time windows and scalp locations (King and Dehaene 2014; Marsicano, Bertini, and Ronconi 2024). Here, we used MVPA to examine the neural correlates of syllable discrimination (here identified as “decoding accuracy”) with and without acoustic exaggeration. The goal was twofold: (1) to determine the extent to which EEG responses could discriminate between the target syllables, and (2) to determine whether acoustically exaggerated syllables could be decoded with greater accuracy than non-exaggerated ones.

The MVPA analysis was implemented using the Amsterdam Decoding and Modeling toolbox (Fahrenfort et al. 2018). A backward decoding model (BDM) was used for first-level (single subject) analysis. BDMs enable the prediction of an experimental variable (in this case, the target syllable /i2/ or /p^{hi}1/) based on an observed pattern of brain activity. The BDM uses linear discriminant analysis to perform decoding (Grootswagers, Wardle, and Carlson 2017). A 5-fold cross-validation procedure was used to train-and-test the classifier. Specifically, the trials were split into five equal-sized folds, with 4 folds as the training set and the remaining fold as the testing set. This process was iterated five times, with each fold being tested once. The classification performance was quantified by area under the curve (AUC). The AUCs from each iteration or fold were averaged to produce a single performance metric for each time point. In this first step, the

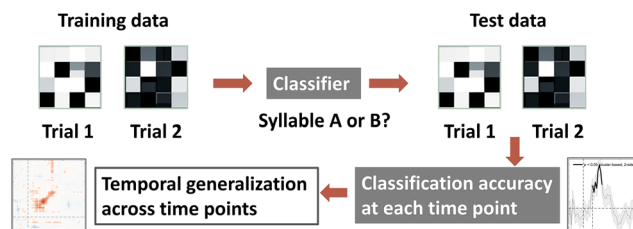


FIGURE 2 | Workflow of the MVPA analyses. The algorithm trains a classifier using EEG response data to predict an experimental variable (here, the target syllable /i2/ or /p^{hi}1/). The classifier is then tested on a separate EEG dataset from the same participant to estimate its classification accuracy, measured by the area under the curve (AUC). Initially, both training and testing are done at the same time point within the EEG trial. Subsequently, the classifier is tested at different times to evaluate its ability to generalize across time, producing a visual representation- the temporal generalization matrix. The extent of color in the temporal generalization matrix indicates the stability of neural processing over time.

classifier was trained and tested at the same time points, which is called diagonal decoding (see Figure 2). Following the subject-level analysis, group-level statistics for MVPA were derived through *t*-tests, which compared the AUC against the chance level of 0.5 accuracy. A *p*-value was produced for each time point as a result of the group-level analysis. Then, a CBPT with 1000 iterations was applied to address multiple comparisons and to identify significant decoding clusters (clusters of time points that significantly distinguish the target syllables).

2.4.4 | MVPA: Temporal Generalization

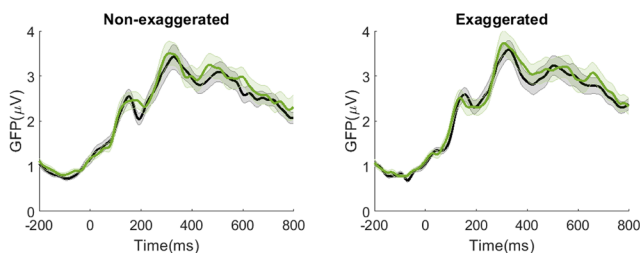
To characterize the temporal dynamics of neural processing, we computed a temporal generalization matrix for each group and condition. The diagonal values in the matrix represent the AUC when training and testing of the classifiers were performed on the same time points, without cross-time validation. These values are equivalent to the diagonal decoding described in the previous section. The off-diagonal values were computed in the same way as the diagonal values, but with the classifiers tested on different time points, indicative of the generalizability of classifiers across time (Figure 2). High off-diagonal values indicate that the neural patterns identified during training are also applicable at other time points. We first analyzed the signals recorded from all channels (excluding the eye channels), and then the left- and right-hemisphere channels separately to gain insights into possible hemispheric lateralization of speech processing.

3 | Results

3.1 | Univariate ERP Analysis Fails to Differentiate Syllables With and Without Acoustic Exaggeration

This section reports the results of conventional waveform analysis. Non-parametric CBPT indicated that in both the non-exaggerated and exaggerated speech conditions, and in both the autistic and TD groups, the GFPs elicited by the /i2/ syllable were not significantly different from those elicited by the /p^{hi}1/ syllable (Figure 3). This lack of syllable difference suggests that

TD: /i2/ vs /p^hi1/



Autism: /i2/ vs /p^hi1/

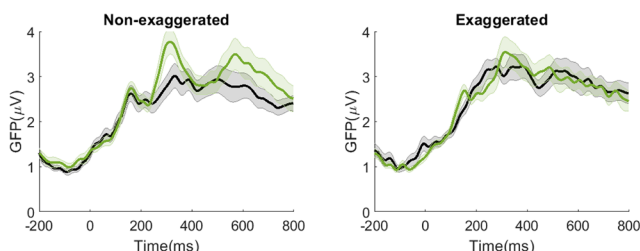


FIGURE 3 | Global-field-power (GFP) waveforms comparing responses to the /i2/ versus /pi1/ syllables in each group. No response difference between syllables was identified, regardless of group or acoustic exaggeration, as evidenced by the cluster-based permutation tests (CBPT).

a univariate ERP measure may not provide neural correlates for syllable discrimination, even in the presence of acoustic exaggeration. Group comparisons revealed no significant difference, regardless of syllable or acoustic exaggeration (Figure S1).

We then examined the effects of acoustic exaggeration on GFP for each syllable using CBPT. In the autism group, but not in the TD group, tonal exaggeration significantly increased the GFP amplitude by around 250 ms ($p=0.019$; Figure 4). No significant GFP differences related to VOT exaggeration were observed in either group.

In the following sections, we directly address whether acoustic exaggeration aided syllable discrimination using the multivariate approach.

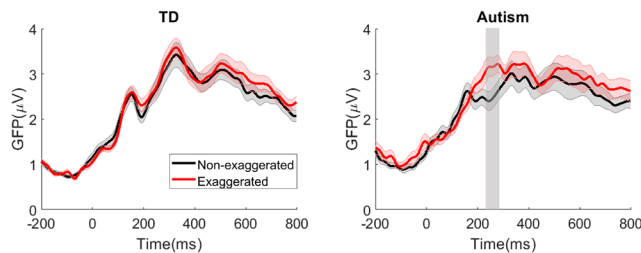
3.2 | MVPA Diagonal Decoding Reveals Enhanced Neural Discrimination of Exaggerated Syllables in Autism Compared With Non-Exaggerated Syllables

This section reports on diagonal decoding, which assesses syllable discrimination with and without acoustic exaggeration. The goals were (1) to determine how well target syllables were decoded from EEG responses and (2) to compare the decoding accuracy of acoustically exaggerated syllables with that of non-exaggerated syllables.

3.2.1 | All Channels

In the non-exaggerated speech condition, neither group showed significant neural decoding of the syllables. For the TD group, CBPT did not identify significant decoding clusters

Syllable /i2/: tone exaggeration



Syllable /p^hi1/: VOT exaggeration

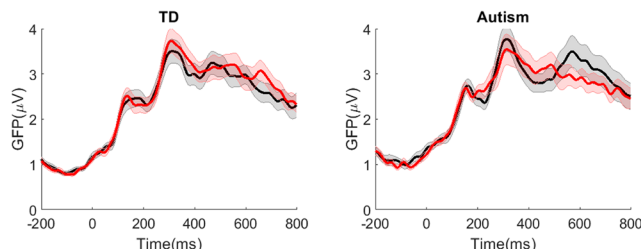


FIGURE 4 | Global-field-power (GFP) waveforms comparing responses with and without acoustic exaggeration for each syllable in each group. The shaded area represents a significant effect of exaggeration on GFP amplitude identified by cluster-based permutation test (CBPT).

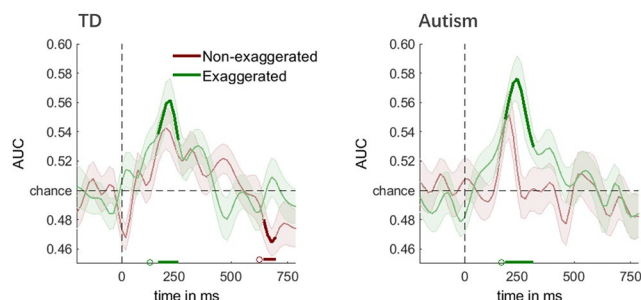


FIGURE 5 | Comparisons of MVPA decoding accuracies for the non-exaggerated and exaggerated speech. Thick lines represent time clusters that are significantly different from chance identified by CBPT.

that distinguished the two syllables (Figure 5). One-sample t -tests were performed on the group-average AUC values, calculated at each time point over the entire trial. This confirmed that their overall decoding accuracies were not significantly different from a chance AUC of 0.5 [$t(54)=1.11$, $p=0.273$, Cohen's $d=0.15$]. Similarly, the autism group showed no significant decoding clusters, and one-sample t -test further showed that their group-average decoding accuracies did not differ from chance [$t(54)=-1.93$, $p=0.059$, Cohen's $d=-0.26$]. Although neither group produced significant decoding clusters, an independent samples t -test for group differences showed that the TD group had a higher overall AUC than the autism group [$t(108)=2.06$, $p=0.042$, Cohen's $d=0.39$] in the non-exaggerated conditions.

In contrast, the exaggerated speech condition elicited significant neural decoding for the syllables in both groups, but with notable differences between them. The TD group showed significant decoding clusters between 172 and 263 ms, peaking at 226 ms ($p=0.006$; Figure 5), according to CBPT analysis. However,

TABLE 3 | Summary of significant comparisons of MVPA diagonal decoding by group.

Comparison	Group	
	Autism	TD
Decoding clusters		
Non-ex syllables	None	None
Ex syllables	190–317 ms***	172–263 ms**
Group-average decoding accuracy versus chance		
Non-ex syllables	NS	NS
Ex syllables	**	NS ($p = 0.054$)
Within-group comparisons: decoding accuracy		
Non-ex versus ex syllables	Ex > Non-ex*	NS
Between-group comparisons: decoding accuracy		
Non-ex syllables	TD > Autism*	
Ex syllables	NS	

Abbreviations: Ex, exaggerated; Non-ex, non-exaggerated; NS, not significant.

* $p < 0.05$.** $p < 0.01$.*** $p < 0.001$.

their group-average decoding accuracy was only approaching a significant difference from chance [$t(54) = 1.97$, $p = 0.054$, Cohen's $d = 0.27$]. The autism group showed significant decoding clusters from 190 to 317 ms, peaking at 244 ms ($p < 0.001$; Figure 5), and their decoding accuracy across the trial window was significantly above chance [$t(54) = 3.14$, $p = 0.003$, Cohen's $d = 0.42$]. When comparing the groups, there was no significant difference in overall decoding accuracy in the exaggerated condition [$t(108) = -1.07$, $p = 0.288$, Cohen's $d = -0.20$].

Furthermore, we used paired sample t -tests to examine the effects of acoustic exaggeration within each group. In the autism group, group-average decoding accuracies for exaggerated syllables were significantly higher than those for non-exaggerated syllables [$t(54) = -5.51$, $p < 0.001$, Cohen's $d = -0.70$]. In contrast, the TD group did not show a significant difference between the exaggerated and non-exaggerated conditions [$t(54) = -0.74$, $p = 0.460$, Cohen's $d = -0.11$].

In sum (Table 3), the decoding cluster results indicate that acoustic exaggeration enhanced the neural discrimination of syllables in the autism and TD groups, an effect that was not detected in the univariate analysis. In addition, t -tests showed that the autism group achieved overall decoding accuracies significantly above the chance for exaggerated syllables, with performance comparable to the TD group.

3.2.2 | Left Versus Right Hemisphere Channels

3.2.2.1 | TD Group. For the non-exaggerated syllables, significant decoding clusters were found only in the right hemisphere electrodes (190–317 ms, $p = 0.006$; 354–499 ms, $p < 0.001$),

with higher overall decoding accuracies than in the left hemisphere [$t(54) = -4.60$, $p < 0.001$, Cohen's $d = -0.53$].

For the exaggerated syllables, significant decoding clusters were found in both hemispheres (left, 117–354 ms peaked at 226 ms, $p < 0.001$; right, 172–299 ms peaked at 226 ms, $p = 0.001$), with the left performing better [$t(54) = 6.16$, $p < 0.001$, Cohen's $d = 0.53$].

3.2.2.2 | Autism Group. For the non-exaggerated syllables, the left hemisphere electrodes showed higher overall decoding accuracies than the right [$t(54) = 6.77$, $p < 0.001$, Cohen's $d = 0.67$], but no significant clusters were identified.

For the exaggerated syllables, the two hemispheres had comparable decoding accuracies [$t(54) = 0.94$, $p = 0.349$, Cohen's $d = 0.07$], with significant clusters in both hemispheres (left, 190–280 ms peaked at 244 ms, $p < 0.001$; right, 135–281 ms peaked at 226 ms, $p < 0.001$).

3.3 | MVPA Temporal Generalization Reveals Distinct Neural Coding Mechanisms in Autism and TD

This section reports the MVPA results in the time domain. The aim was to determine whether the neural pattern supporting above-chance syllable discrimination is temporally stable. The temporal generalization analysis was performed only for the exaggerated syllable conditions, as the non-exaggerated syllables were undecodable—did not produce neural indicators of discrimination.

3.3.1 | All Channels

The temporal generalization matrices (Figure 6) showed that the autism group had a predominantly diagonal decoding pattern, indicating limited generalization across time. In contrast, the TD group displayed a square, off-diagonal pattern in the early decoding window.

To quantify this, we tested the classifier trained at the AUC peak across all time points. In the TD group, the classifier at peak AUC (226 ms) generalized to 117–281 ms ($p = 0.008$). In the autism group, the classifier at peak AUC (244 ms) generalized to a later time window, 208–281 ms ($p = 0.006$). These results suggest that decoding is earlier and more temporally reliable in the TD group compared with the autism group.

3.3.2 | Left Versus Right Hemisphere Channels

In the TD group, the left hemisphere showed earlier and more extensive generalization than the right (Figure 7). Further analyses at peak AUCs for each hemisphere (left peak: 226 ms; right peak: 226 ms) verified this pattern, showing earlier generalization on the left than on the right (left, 81–263 ms, $p = 0.012$; right, 172–281 ms, $p = 0.010$).

In the autism group, the two hemispheres showed similar diagonal patterns of decoding (Figure 7). Further analyses with

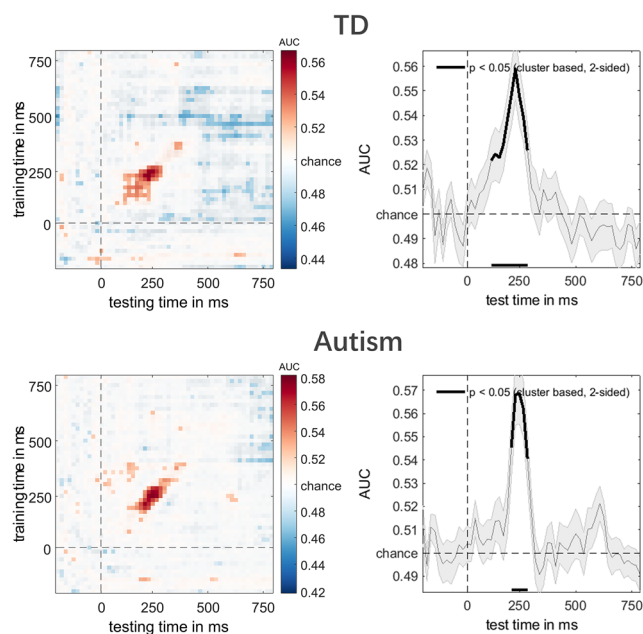


FIGURE 6 | Temporal generalization matrices computed using all channels; waveforms on the right depict generalization when tested with classifier at peak AUC in each group.

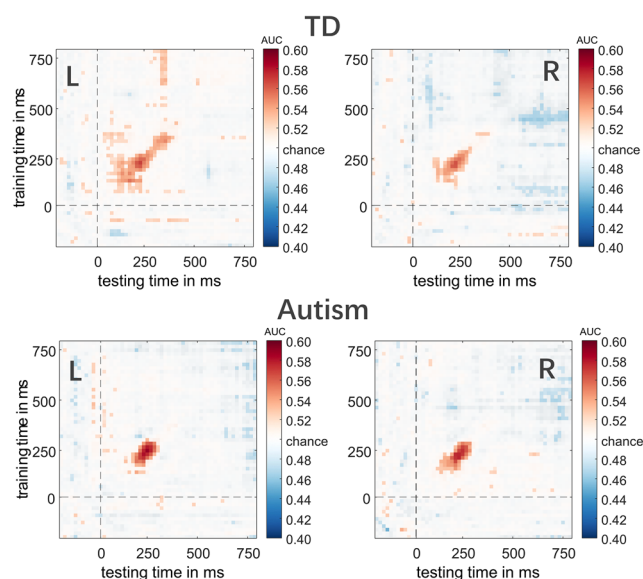


FIGURE 7 | Temporal generalization matrices of the left and right hemisphere channels.

training classifiers at peak AUCs (left peak: 244 ms; right peak: 226 ms) identified bilateral generalization beyond 200 ms (left, 208–299 ms, $p = 0.006$; right, 208–281 ms, $p = 0.022$).

Pearson analyses between MVPA measures and age, MA, SRS scores, CARS scores, and ABC scores did not reveal meaningful correlations.

4 | Discussion

The present study investigates the impact of acoustic exaggeration on speech processing of autistic children with speech

delay. This represents the first attempt to apply a multivariate approach with EEG in the domain of speech and language perception in autism. The results demonstrated that acoustic exaggeration improved the neural processing of speech sounds in the autism group, as evidenced by the elevated MVPA classification/decoding performance for acoustically exaggerated syllables compared with those for non-exaggerated syllables. A time-resolved analysis further revealed that the effects of acoustic exaggeration in autism exhibited distinct temporal dynamics with regard to hemispheric lateralization, suggesting differential speech processing mechanisms.

4.1 | Acoustic Exaggeration Improves Autistic Speech Discrimination

A main finding of the current study is that acoustic exaggeration at the phonemic level significantly influenced the neural correlates of syllable discrimination in autistic children. It provides additional, or at least comparable, benefit for the autistic children in comparison with their TD peers, a phenomenon that would not be predicted based on past literature demonstrating limited attentional orientation or cortical responsivity to CDS in autism (Chen et al. 2021; Kuhl et al. 2005; Paul et al. 2007; Pierce et al. 2023).

Research using naturalistic and continuous CDS (motherese) has reported a reduced level of attention to CDS during the first years of life, which has been identified as a putative pre-symptomatic marker for autism (Pierce et al. 2023) and an impedance for social and language development (Watson et al. 2012). Along this line, Chen et al. (2021) found that 4–11-year-old autistic children displayed a diminished cortical response enhancement in the early P1 window to formant-exaggerated vowels relative to their TD peers. The lack of enhancement was considered to hinder the phonological acquisition of autistic children. The results of their study appear to be contrasting with ours, which showed enhanced GFP associated with tonal exaggeration in autism, but not in TD. This discrepancy could be attributed to differences in the sample age, the specific phonemic features investigated (formant space vs. lexical tone), or analytical techniques used in the ERP analysis. More importantly, this previous study did not sufficiently address the impact of acoustic exaggeration on speech discrimination outcomes.

The current study presents the first evidence that selective acoustic exaggeration of phonemic features enhances the accuracy of syllable discrimination in young autistic listeners, as indicated by neural decoding correlates. The auditory stimuli used in the current study differ considerably from naturalistically recorded CDS. The only acoustic exaggerations implemented here were either VOT or F0 range, which are, respectively, the deterministic features of the target stop consonant and lexical tone. In consequence, the various canonical features of CDS including pitch height, intensity, and formant-frequency dispersion (Hilton et al. 2022) remained unchanged. This rigorous control of the acoustic variables eliminated the potential influence of social affect and other phonemically irrelevant factors, thus allowing the attribution of the effects on decoding accuracy to the acoustic exaggeration alone.

The acoustic quality of the exaggerated speech used in this study may find a correspondence in the non-social language material, which is known as frequently sourcing autistics' initial language productions (TVs, radios, etc.; Kissine 2021; Mottron, Ostrolenk, and Gagnon 2021). Unlike the acoustic variations inherent to naturalistic, socially rich CDS or motherese, the acoustically exaggerated speech used here has minimalistic acoustic differences from non-exaggerated speech, with the exception of the key phonemic variables, namely, the VOT of the stop consonant and the F0 range of the lexical tone. Such abstracted acoustic exaggeration serves to accentuate phonemic contrasts while excluding irrelevant acoustic variations. Previous study with autistic individuals has shown that an increased neural response for a vowel contrast is indeed disturbed by irrelevant pitch variations, potentially reflecting a diminished capacity to extract invariant phonemic categories from the highly variable speech input required for language acquisition (Kujala, Lepistö, and Näätänen 2013; Lepistö et al. 2008). In the context of lower oscillatory synchrony or "neural jitter" in EEG signals during speech processing in autistic children (Yu et al. 2018), *targeted* acoustic exaggerations may prove to be a more effective approach in supporting cortical speech processing in autism.

4.2 | Distinct Lateralization of Temporal Dynamics Underlying Speech Processing Between Autism and TD

In the TD group, neural decoding was predominantly lateralized to the left hemisphere, with the exception of diagonal decoding in the non-exaggerated condition. The finding that the right, but not the left hemisphere, significantly decoded the non-exaggerated syllables in the TD group is in alignment with existing literature that suggests the role of the right hemisphere in speech perception involving frequency modulations (e.g., lexical tone contrast; Luo et al. 2006) or, in some cases, VOT differences (Molfese 1998). However, when the phonemic features were acoustically exaggerated, decoding in the TD group became left-lateralized for both diagonal decoding and temporal generalization. This shift of lateralization may be due to the enhanced prominence of speech features that potentially triggered the specialized network in the left hemisphere of TD. In comparison, the autism group exhibited a bilateral decoding pattern, irrespective of acoustic exaggeration.

The presence and nature of temporal generalization are indicative of the stability of neural coding over time (King and Dehaene 2014). The TD group showed a left lateralization and an early pattern of temporal generalization starting within 100 ms. In contrast, the autism group displayed a bilateral, relatively late, and narrow temporal generalization (diagonal pattern in the generalization matrices). The TD results may reflect a more rapid and reliable speech encoding system for syllable material, driven by left-hemisphere functions. This is consistent with the established differences in the computational properties of the two hemispheres (Poeppel, Idsardi, and van Wassenhove 2008; Zatorre 2022). The left hemisphere is known to preferentially represent fast-changing and temporal properties in the auditory signal (e.g., VOT), and the right hemisphere is more attuned to slow-changing and spectral properties (e.g., F0 contour; Boemio et al. 2005; Luo et al. 2006; Zaehle et al. 2004; but

see Molfese 1998; Simos, Molfese, and Brenden 1997 for VOT processing in the right hemisphere). The leftward and rapid temporal generalization in the TD group, despite the smaller exaggeration for the VOT than that for the F0 range (1.19 vs. 1.95), suggests a more efficient syllable detection using the initial consonant cue, which may involve a specialized brain network.

In contrast, the lack of leftward asymmetry in the autistic group may indicate an absence or delayed development of functional specialization toward the left hemisphere (Haesen, Boets, and Wagemans 2011; Li et al. 2023; Lindell and Hudry 2013; Yu and Wang 2021). It is possible that autistic children may rely less on transient and rapid acoustic cues, such as VOT, which are processed predominantly by the left hemisphere. Notably, we observed a tonal exaggeration effect on the GFP amplitude only in the autism group (Figure 4). This autism-specific effect, in conjunction with the MVPA results, leads us to speculate a pitch-based strategy in their speech encoding system (e.g., Wang and Xu 2024). This interpretation is consistent with the enhanced pitch discrimination evident in older autistic people (Bonnell et al. 2003).

Taken together, selective acoustic exaggerations of stop consonant and lexical tone features result in an immediate improvement in neural decoding of syllables in both autistic and TD children. Nevertheless, the relative weight of VOT and tonal cues may differ between the two groups. It may be the case that autistic children require a greater extent of VOT exaggeration for syllable detection in comparison to their TD counterparts.

4.3 | Limitations

The study has several limitations. The first is the relatively broad age range of the participants, which may have introduced variability in developmental stages affecting auditory functions and responsiveness to CDS. Although this age distribution is narrower than that in previous studies (e.g., Chen et al. 2021), it emphasizes the necessity for future research to minimize developmental variability and to better isolate the effects of CDS and acoustic exaggeration. The second limitation is the specificity of the participant group. The data were collected exclusively from autistic children with speech delays. These children may represent a specific subgroup within the autism spectrum with distinct auditory processing and language acquisition mechanisms (Chiodo, Mottron, and Majerus 2019; Mottron, Ostrolenk, and Gagnon 2021). As a result, our findings may not be generalizable to all autistic children with different language development profiles (e.g., Bonnell et al. 2010). Third, while autistic children had significantly lower mental age, as measured by receptive language ability, it was not possible to isolate the potential effect of this difference from the results. However, the use of a passive listening paradigm minimized its impact by removing the need for active cognitive or linguistic engagement. Moreover, language differences may be intrinsic to the autistic phenotype, which makes it difficult to disentangle these effects from more specific information processing differences (e.g., Schaeffer et al. 2023). Finally, the current study used a single stimulus type—syllables—to assess the impact of exaggerated acoustic features on speech processing. While syllables are fundamental units of speech, the results might not be generalizable to higher-order

linguistic structures, which limits the study's scope and ecological validity.

4.4 | Implications and Future Directions

The present study investigates the differential effects of speech acoustics on autistic and typical children's speech discrimination. Exaggerated speech provides amplified acoustic features to enhance the clarity of auditory signal, thereby increasing the contrast and discriminability of different speech sounds. The findings have implications for sensory-based strategies aimed at supporting speech communication and learning in young autistic children. For example, at the beginning of speech and language training for autistic children, it might be useful to consider non-social-based materials with targeted acoustic exaggerations to leverage autistic children's proneness for benefiting from such signals. How this finding can be transferred into measurable progress in speech encoding and use remains to be studied.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (31900775) and the Plan on enhancing scientific research in Guangzhou Medical University (2024SRP123). We thank our participants and their families for their contribution to this research, and the involved staff and students for their assistance with data collection.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are openly available in OSF at https://osf.io/b6qe2/?view_only=c9ef4d65578149438ae952b152e1221f.

References

- Abrams, D. A., T. Nicol, S. Zecker, and N. Kraus. 2008. "Right-Hemisphere Auditory Cortex Is Dominant for Coding Syllable Patterns in Speech." *Journal of Neuroscience* 28, no. 15: 3958–3965. <https://doi.org/10.1523/JNEUROSCI.0187-08.2008>.
- American Psychiatric Association. 2013. *Diagnostic and Statistical Manual of Mental Disorders*. 5th ed. Washington, DC: American Psychiatric Association.
- Boemio, A., S. Fromm, A. Braun, and D. Poeppel. 2005. "Hierarchical and Asymmetric Temporal Sensitivity in Human Auditory Cortices." *Nature Neuroscience* 8, no. 3: 389–395. <https://doi.org/10.1038/nn1409>.
- Bonnel, A., S. McAdams, B. Smith, et al. 2010. "Enhanced Pure-Tone Pitch Discrimination Among Persons With Autism But Not Asperger Syndrome." *Neuropsychologia* 48, no. 9: 2465–2475. <https://doi.org/10.1016/j.neuropsychologia.2010.04.020>.
- Bonnel, A., L. Mottron, I. Peretz, M. Trudel, E. Gallun, and A. Bonnel. 2003. "Enhanced Pitch Sensitivity in Individuals With Autism: A Signal Detection Analysis." *Journal of Cognitive Neuroscience* 15, no. 2: 226–235.
- Burnham, E., J. L. Gamache, T. Bergeson, and L. Dilley. 2013. "Voice-Onset Time in Infant-Directed Speech Over the First Year and a Half." *Proceedings of Meetings on Acoustics* 19, no. 1: 226.

- Cen, C.-Q., Y.-Y. Liang, Q.-R. Chen, et al. 2007. "Investigating the Validation of the Chinese Mandarin Version of the Social Responsiveness Scale in a Mainland China Child Population." *BMC Psychiatry* 17, no. 1. <https://doi.org/10.1186/s12888-016-1185-y>.
- Chen, F., and G. Peng. 2021. "Categorical Perception of Pitch Contours and Voice Onset Time in Mandarin-Speaking Adolescents With Autism Spectrum Disorders." *Journal of Speech, Language, and Hearing Research* 64, no. 11: 4468–4484. https://doi.org/10.1044/2021_JSLHR-20-00725.
- Chen, F., H. Zhang, H. Ding, S. Wang, G. Peng, and Y. Zhang. 2021. "Neural Coding of Formant-Exaggerated Speech and Nonspeech in Children With and Without Autism Spectrum Disorders." *Autism Research* 14, no. 7: 1357–1374. <https://doi.org/10.1002/aur.2509>.
- Chen, Y., E. Tang, H. Ding, and Y. Zhang. 2022. "Auditory Pitch Perception in Autism Spectrum Disorder: A Systematic Review and Meta-Analysis." *Journal of Speech, Language, and Hearing Research* 65, no. 12: 4866–4886. https://doi.org/10.1044/2022_JSLHR-22-00254.
- Chiodo, L., L. Mottron, and S. Majerus. 2019. "Preservation of Categorical Perception for Speech in Autism With and Without Speech Onset Delay." *Autism Research* 12, no. 11: 1609–1622. <https://doi.org/10.1002/aur.2134>.
- Constantino, J. N., and C. P. Gruber. 2009. "Social Responsiveness Scale." *Western Psychological Services*.
- Cooper, R. P., and R. N. Aslin. 1990. "Preference for Infant-Directed Speech in the First Month After Birth." *Child Development* 61, no. 5: 1584–1595.
- Cox, C., C. Bergmann, E. Fowler, et al. 2023. "A Systematic Review and Bayesian Meta-Analysis of the Acoustic Features of Infant-Directed Speech." *Nature Human Behaviour* 7, no. 1: 114–133. <https://doi.org/10.1038/s41562-022-01452-1>.
- Cristia, A. 2010. "Phonetic Enhancement of Sibilants in Infant-Directed Speech." *Journal of the Acoustical Society of America* 128, no. 1: 424–434. <https://doi.org/10.1121/1.3436529>.
- Deemyad, T. 2022. "Lateralized Changes in Language Associated Auditory and Somatosensory Cortices in Autism." *Frontiers in Systems Neuroscience* 16: 787448. <https://doi.org/10.3389/fnsys.2022.787448>.
- Delorme, A., and S. Makeig. 2004. "EEGLAB: An Open Source Toolbox for Analysis of Single-Trial EEG Dynamics Including Independent Component Analysis." *Journal of Neuroscience Methods* 134, no. 1: 9–21. <https://doi.org/10.1016/j.jneumeth.2003.10.009>.
- D'Oodorico, L., and V. Jacob. 2006. "Prosodic and Lexical Aspects of Maternal Linguistic Input to Late-Talking Toddlers." *International Journal of Language and Communication Disorders* 41, no. 3: 293–311.
- Droucker, D., S. Curtin, and A. Vouloumanos. 2012. "Linking Infant-Directed Speech and Face Preferences to Language Outcomes in Infants at Risk for Autism Spectrum Disorder." *Journal of Speech, Language, and Hearing Research* 56, no. 2: 567–576. <https://doi.org/10.1044/1092-4388.2012.11-0266>.
- Dunst, C. J., A. Simkus, and D. W. Hamby. 2012. "Preference for Infant-Directed Speech in Preverbal Young Children." *Center for Early Literacy Learning* 5, no. 3: 1–10.
- Edgar, J. C., S. Y. Khan, L. Blaskey, et al. 2015. "Neuromagnetic Oscillations Predict Evoked-Response Latency Delays and Core Language Deficits in Autism Spectrum Disorders." *Journal of Autism and Developmental Disorders* 45, no. 2: 395–405. <https://doi.org/10.1007/s10803-013-1904-x>.
- Englund, K. T. 2005. "Voice Onset Time in Infant Directed Speech Over the First Six Months." *First Language* 25, no. 2: 219–234. <https://doi.org/10.1177/0142723705050286>.
- Fahrenfort, J. J., J. van Driel, S. van Gaal, and C. N. L. Olivers. 2018. "From ERPs to MVPA Using the Amsterdam Decoding and Modeling

- Toolbox (ADAM)." *Frontiers in Neuroscience* 12: 368. <https://doi.org/10.3389/fnins.2018.00368>.
- Fernald, A., and T. Simon. 1984. "Expanded Intonation Contours in Mothers' Speech to Newborns." *Developmental Psychology* 20, no. 1: 104–113.
- Finch, K. H., A. M. Seery, M. R. Talbott, C. A. Nelson, and H. Tager-Flusberg. 2017. "Lateralization of ERPs to Speech and Handedness in the Early Development of Autism Spectrum Disorder." *Journal of Neurodevelopmental Disorders* 9: 4. <https://doi.org/10.1186/s11689-017-9185-x>.
- Fish, M. S., A. García-Sierra, N. Ramírez-Esparza, and P. K. Kuhl. 2017. "Infant-Directed Speech in English and Spanish: Assessments of Monolingual and Bilingual Caregiver VOT." *Journal of Phonetics* 63: 19–34. <https://doi.org/10.1016/j.wocn.2017.04.003>.
- Fisher, C., and H. Tokura. 1995. "The Give-New Contract in Speech to Infants."
- Foss-Feig, J. H., L. D. Kwakye, C. J. Cascio, et al. 2010. "An Extended Multisensory Temporal Binding Window in Autism Spectrum Disorders." *Experimental Brain Research* 203, no. 2: 381–389. <https://doi.org/10.1007/s00221-010-2240-4>.
- Gandal, M. J., J. C. Edgar, R. S. Ehrlichman, M. Mehta, T. P. Roberts, and S. J. Siegel. 2010. "Validating Gamma Oscillations and Delayed Auditory Responses as Translational Biomarkers of Autism." *Biological Psychiatry* 68, no. 12: 1100–1106. <https://doi.org/10.1016/j.biopsych.2010.09.031>.
- Gangopadhyay, I., and M. Kaushanskaya. 2020. "The Role of Speaker Eye Gaze and Mutual Exclusivity in Novel Word Learning by Monolingual and Bilingual Children." *Journal of Experimental Child Psychology* 197: 104878. <https://doi.org/10.1016/j.jecp.2020.104878>.
- Gong, Z., and D. Guo. 1984. "An Intelligence Screening Test for Preschool and Primary School Children—Picture Vocabulary Test." *Acta Psychologica Sinica* 16, no. 4: 46.
- Gratier, M., and E. Devouche. 2011. "Imitation and Repetition of Prosodic Contour in Vocal Interaction at 3 Months." *Developmental Psychology* 47, no. 1: 67–76. <https://doi.org/10.1037/a0020722>.
- Grootswagers, T., S. G. Wardle, and T. A. Carlson. 2017. "Decoding Dynamic Brain Patterns From Evoked Responses: A Tutorial on Multivariate Pattern Analysis Applied to Time Series Neuroimaging Data." *Journal of Cognitive Neuroscience* 29, no. 4: 677–697. https://doi.org/10.1162/jocn_a_01068.
- Haesen, B., B. Boets, and J. Wagemans. 2011. "A Review of Behavioural and Electrophysiological Studies on Auditory Processing and Speech Perception in Autism Spectrum Disorders." *Research in Autism Spectrum Disorders* 5, no. 2: 701–714. <https://doi.org/10.1016/j.rasd.2010.11.006>.
- Hartman, K. M., N. B. Ratner, and R. S. Newman. 2017. "Infant-Directed Speech (IDS) Vowel Clarity and Child Language Outcomes." *Journal of Child Language* 44, no. 5: 1140–1162. <https://doi.org/10.1017/S0305000916000520>.
- Hickok, G., and D. Poeppel. 2007. "The Cortical Organization of Speech Processing [Research Support, N.I.H., Extramural Review]." *Nature Reviews* 8, no. 5: 393–402. <https://doi.org/10.1038/nrn2113>.
- Hilton, C. B., C. J. Moser, M. Bertolo, et al. 2022. "Acoustic Regularities in Infant-Directed Speech and Song Across Cultures." *Nature Human Behaviour* 6, no. 11: 1545–1556. <https://doi.org/10.1038/s41562-022-01410-x>.
- Huang, D., L. Yu, X. Wang, Y. Fan, S. Wang, and Y. Zhang. 2018. "Distinct Patterns of Discrimination and Orienting for Temporal Processing of Speech and Nonspeech in Chinese Children With Autism: An Event-Related Potential Study." *European Journal of Neuroscience* 47, no. 6: 662–668. <https://doi.org/10.1111/ejn.13657>.
- Kalashnikova, M., and D. Burnham. 2018. "Infant-Directed Speech From Seven to Nineteen Months Has Similar Acoustic Properties But Different Functions." *Journal of Child Language* 45, no. 5: 1035–1053. <https://doi.org/10.1017/S0305000917000629>.
- Kasai, K., O. Hashimoto, Y. Kawakubo, et al. 2005. "Delayed Automatic Detection of Change in Speech Sounds in Adults With Autism: A Magnetoencephalographic Study." *Clinical Neurophysiology* 116, no. 7: 1655–1664. <https://doi.org/10.1016/j.clinph.2005.03.007>.
- King, J. R., and S. Dehaene. 2014. "Characterizing the Dynamics of Mental Representations: The Temporal Generalization Method." *Trends in Cognitive Sciences* 18, no. 4: 203–210. <https://doi.org/10.1016/j.tics.2014.01.002>.
- Kissine, M. 2021. "Autism, Constructionism, and Nativism." *Language* 97, no. 3: e139–e160. <https://doi.org/10.1353/lan.2021.0055>.
- Kuhl, P. K. 2010. "Brain Mechanisms in Early Language Acquisition." *Neuron* 67, no. 5: 713–727. <https://doi.org/10.1016/j.neuron.2010.08.038>.
- Kuhl, P. K., J. E. Andruski, I. A. Chistovich, et al. 1997. "Cross-Language Analysis of Phonetic Units in Language Addressed to Infants." *Science* 277, no. 5326: 684–686. <https://doi.org/10.1126/science.277.5326.684>.
- Kuhl, P. K., S. Coffey-Corina, D. Padden, and G. Dawson. 2005. "Links Between Social and Linguistic Processing of Speech in Preschool Children With Autism: Behavioral and Electrophysiological Measures." *Developmental Science* 8, no. 1: F1–F12.
- Kujala, T., T. Lepistö, and R. Näätänen. 2013. "The Neural Basis of Aberrant Speech and Audition in Autism Spectrum Disorders." *Neuroscience and Biobehavioral Reviews* 37, no. 4: 697–704. <https://doi.org/10.1016/j.neubiorev.2013.01.006>.
- Kwakye, L. D., J. H. Foss-Feig, C. J. Cascio, W. L. Stone, and M. T. Wallace. 2011. "Altered Auditory and Multisensory Temporal Processing in Autism Spectrum Disorders." *Frontiers in Integrative Neuroscience* 4: 129. <https://doi.org/10.3389/fnint.2010.00129>.
- Lai, B., A. Yi, F. Zhang, et al. 2024. "Atypical Brain Lateralization for Speech Processing at the Sublexical Level in Autistic Children Revealed by fNIRS." *Scientific Reports* 14, no. 1: 2776. <https://doi.org/10.1038/s41598-024-53128-7>.
- Lehmann, D., and W. Skrandies. 1984. "Spatial Analysis of Evoked Potentials in Man—A Review." *Progress in Neurobiology* 23, no. 3: 227–250.
- Lepistö, T., M. Kajander, R. Vanhala, et al. 2008. "The Perception of Invariant Speech Features in Children With Autism." *Biological Psychology* 77, no. 1: 25–31. <https://doi.org/10.1016/j.biopsycho.2007.08.010>.
- Li, Q., W. Zhao, L. Palaniyappan, and S. Guo. 2023. "Atypical Hemispheric Lateralization of Brain Function and Structure in Autism: A Comprehensive Meta-Analysis Study." *Psychological Medicine* 1–12: 6702–6713. <https://doi.org/10.1017/S0033291723000181>.
- Li, W. Q., X. Liu, Y. Dai, and Q. Cheng. 2018. "Age of Diagnosis of Autism Spectrum Disorder in Children and Factors Influencing the Age of Diagnosis." *Chinese Journal of Contemporary Pediatrics* 20, no. 10: 799–803. <https://doi.org/10.7499/j.issn.1008-8830.2018.10.003>.
- Lindell, A. K., and K. Hudry. 2013. "Atypicalities in Cortical Structure, Handedness, and Functional Lateralization for Language in Autism Spectrum Disorders." *Neuropsychology Review* 23, no. 3: 257–270. <https://doi.org/10.1007/s11065-013-9234-5>.
- Liu, H., P. K. Kuhl, and F. M. Tsao. 2003. "An Association Between Mothers' Speech Clarity and Infants' Mothers' Speech Clarity and Infants' Discrimination Speech Discrimination Skills." *Developmental Science* 6, no. 3: F1–F10.
- Luo, H., J. T. Ni, Z. H. Li, et al. 2006. "Opposite Patterns of Hemisphere Dominance for Early Auditory Processing of Lexical Tones and Consonants." *Proceedings of the National Academy of Sciences of the*

- United States of America 103, no. 51: 19558–19563. <https://doi.org/10.1073/pnas.0607065104>.
- Lyakso, E. E., O. V. Frolova, and A. S. Grigorev. 2014. "Infant Vocalizations at the First Year of Life Predict Speech Development at 2–7 Years: Longitudinal Study." *Psychology* 5, no. 12: 1433–1445. <https://doi.org/10.4236/psych.2014.512154>.
- Maris, E., and R. Oostenveld. 2007. "Nonparametric Statistical Testing of EEG- and MEG-Data." *Journal of Neuroscience Methods* 164, no. 1: 177–190. <https://doi.org/10.1016/j.jneumeth.2007.03.024>.
- Marsicano, G., C. Bertini, and L. Ronconi. 2024. "Decoding Cognition in Neurodevelopmental, Psychiatric and Neurological Conditions With Multivariate Pattern Analysis of EEG Data." *Neuroscience and Biobehavioral Reviews* 164: 105795. <https://doi.org/10.1016/j.neubiorev.2024.105795>.
- Miron, O., R. E. Delgado, C. F. Delgado, et al. 2021. "Prolonged Auditory Brainstem Response in Universal Hearing Screening of Newborns With Autism Spectrum Disorder." *Autism Research* 14, no. 1: 46–52. <https://doi.org/10.1002/aur.2422>.
- Molfese, D. L. 1998. "Electrophysiological Correlates of Early Speech Perception and Language Development During Infancy and Early Childhood." *Advances in Psychology* 135: 107–153.
- Morrel, J., K. Singapuri, R. J. Landa, and R. Reetzke. 2023. "Neural Correlates and Predictors of Speech and Language Development in Infants at Elevated Likelihood for Autism: A Systematic Review." *Frontiers in Human Neuroscience* 17: 1211676. <https://doi.org/10.3389/fnhum.2023.1211676>.
- Mottron, L., A. Ostrolenk, and D. Gagnon. 2021. "In Prototypical Autism, the Genetic Ability to Learn Language Is Triggered by Structured Information, Not Only by Exposure to Oral Language." *Genes* 12, no. 8: 112. <https://doi.org/10.3390/genes12081112>.
- Narayan, C. R., and L. C. McDermott. 2016. "Speech Rate and Pitch Characteristics of Infant-Directed Speech: Longitudinal and Cross-Linguistic Observations." *Journal of the Acoustical Society of America* 139, no. 3: 1272–1281. <https://doi.org/10.1121/1.4944634>.
- Norrelgen, F., E. Fernell, M. Eriksson, et al. 2015. "Children With Autism Spectrum Disorders Who Do Not Develop Phrase Speech in the Preschool Years." *Autism* 19, no. 8: 934–943. <https://doi.org/10.1177/1362361314556782>.
- O'Connor, K. 2012. "Auditory Processing in Autism Spectrum Disorder: A Review." *Neuroscience and Biobehavioral Reviews* 36, no. 2: 836–854. <https://doi.org/10.1016/j.neubiorev.2011.11.008>.
- Paul, R., K. Chawarska, C. Fowler, D. Cicchetti, and F. Volkmar. 2007. "'Listen My Children and You Shall Hear': Auditory Preferences in Toddlers With Autism Spectrum Disorders." *Journal of Speech, Language, and Hearing Research* 50, no. 5: 1350–1364. [https://doi.org/10.1044/1092-4388\(2007/094\)](https://doi.org/10.1044/1092-4388(2007/094)).
- Pierce, K., V. H. Gazestani, E. Bacon, et al. 2019. "Evaluation of the Diagnostic Stability of the Early Autism Spectrum Disorder Phenotype in the General Population Starting at 12 Months." *JAMA Pediatrics* 173: 578–587. <https://doi.org/10.1001/jamapediatrics.2019.0624>.
- Pierce, K., T. H. Wen, J. Zahiri, et al. 2023. "Level of Attention to Motherese Speech as an Early Marker of Autism Spectrum Disorder." *JAMA Network Open* 6, no. 2: e2255125. <https://doi.org/10.1001/jamanetworkopen.2022.55125>.
- Poeppel, D., W. J. Idsardi, and V. van Wassenhove. 2008. "Speech Perception at the Interface of Neurobiology and Linguistics." *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 363, no. 1493: 1071–1086. <https://doi.org/10.1098/rstb.2007.2160>.
- Samson, F., K. L. Hyde, A. Bertone, et al. 2011. "Atypical Processing of Auditory Temporal Complexity in Autistics." *Neuropsychologia* 49, no. 3: 546–555. <https://doi.org/10.1016/j.neuropsychologia.2010.12.033>.
- Sang, B., and X. Liao. 1990. "The Revision of Trail Norm of Peabody Picture Vocabulary Test Revised (PPVT-R) in Shanghai Proper." *Psychological Science* 5: 3.
- Schaeffer, J., M. Abd El-Raziq, E. Castroviejo, et al. 2023. "Language in Autism: Domains, Profiles and Co-Occurring Conditions." *Journal of Neural Transmission (Vienna)* 130, no. 3: 433–457. <https://doi.org/10.1007/s00702-023-02592-y>.
- Senju, A., and G. Csibra. 2008. "Gaze Following in Human Infants Depends on Communicative Signals." *Current Biology* 18, no. 9: 668–671. <https://doi.org/10.1016/j.cub.2008.03.059>.
- Simos, P. G., D. L. Molfese, and R. A. Brenden. 1997. "Behavioral and Electrophysiological Indices of Voicing-Cue Discrimination: Laterality Patterns and Development." *Brain and Language* 57, no. 1: 122–150. <https://doi.org/10.1006/brln.1997.1836>.
- Spinelli, M., M. Fasolo, and J. Mesman. 2017. "Does Prosody Make the Difference? A Meta-Analysis on Relations Between Prosodic Aspects of Infant-Directed Speech and Infant Outcomes." *Developmental Review* 44: 1–18. <https://doi.org/10.1016/j.dr.2016.12.001>.
- Steinschneider, M., I. O. Volkov, Y. I. Fishman, H. Oya, J. C. Arezzo, and M. A. Howard 3rd. 2005. "Intracortical Responses in Human and Monkey Primary Auditory Cortex Support a Temporal Processing Mechanism for Encoding of the Voice Onset Time Phonetic Parameter." *Cerebral Cortex* 15, no. 2: 170–186. <https://doi.org/10.1093/cercor/bhh120>.
- Tager-Flusberg, H., R. Paul, and C. Lord. 2005. "Language and Communication in Autism." *Handbook of Autism and Pervasive Developmental Disorders* 1: 335–364.
- Tallal, P., S. Miller, and R. H. Fitch. 1993. "Neurobiological Basis of Speech: A Case for the Preeminence of Temporal Processing." *Annals of the New York Academy of Sciences* 682: 27–47. <https://doi.org/10.1111/j.1749-6632.1993.tb22957.x>.
- Tang, P., N. Xu Rattanasone, I. Yuen, and K. Demuth. 2017. "Phonetic Enhancement of Mandarin Vowels and Tones: Infant-Directed Speech and Lombard Speech." *Journal of the Acoustical Society of America* 142, no. 2: 493–503. <https://doi.org/10.1121/1.4995998>.
- Tardif, T., and P. Fletcher. 2008. "Chinese Communicative Development Inventories: User's Guide and Manual."
- Wang, T., and M. Xu. 2024. *Lexical Tone Perception and Comprehension in Mandarin-Speaking Children With Autism Spectrum Disorder*. Leiden, the Netherlands: Speech Prosody.
- Wang, X., S. Wang, Y. Fan, D. Huang, and Y. Zhang. 2017. "Speech-Specific Categorical Perception Deficit in Autism: An Event-Related Potential Study of Lexical Tone Processing in Mandarin-Speaking Children." *Scientific Reports* 7, no. 1: 43254. <https://doi.org/10.1038/srep43254>.
- Watson, L. R., J. E. Roberts, G. T. Baranek, K. C. Mandulak, and J. C. Dalton. 2012. "Behavioral and Physiological Responses to Child-Directed Speech of Children With Autism Spectrum Disorders or Typical Development." *Journal of Autism and Developmental Disorders* 42, no. 8: 1616–1629. <https://doi.org/10.1007/s10803-011-1401-z>.
- Werker, J. F., and T. K. Hensch. 2015. "Critical Periods in Speech Perception: New Directions." *Annual Review of Psychology* 66: 173–196. <https://doi.org/10.1146/annurev-psych-010814-015104>.
- Yu, L., Y. Fan, Z. Deng, D. Huang, S. Wang, and Y. Zhang. 2015. "Pitch Processing in Tonal-Language-Speaking Children With Autism: An Event-Related Potential Study." *Journal of Autism and Developmental Disorders* 45, no. 11: 3656–3667. <https://doi.org/10.1007/s10803-015-2510-x>.
- Yu, L., D. Huang, S. Wang, X. Wu, Y. Chen, and Y. Zhang. 2021. "Evidence of Altered Cortical Processing of Dynamic Lexical Tone

- Pitch Contour in Chinese Children With Autism.” *Neuroscience Bulletin* 37, no. 11: 1605–1608. <https://doi.org/10.1007/s12264-021-00752-2>.
- Yu, L., and S. Wang. 2021. “Aberrant Auditory System and Its Developmental Implications for Autism.” *Science China Life Sciences* 64, no. 6: 861–878. <https://doi.org/10.1007/s11427-020-1863-6>.
- Yu, L., S. Wang, D. Huang, X. Wu, and Y. Zhang. 2018. “Role of Inter-Trial Phase Coherence in Atypical Auditory Evoked Potentials to Speech and Nonspeech Stimuli in Children With Autism.” *Clinical Neurophysiology* 129, no. 7: 1374–1382. <https://doi.org/10.1016/j.clinph.2018.04.599>.
- Zaehle, T., T. Wustenberg, M. Meyer, and L. Jancke. 2004. “Evidence for Rapid Auditory Perception as the Foundation of Speech Processing: A Sparse Temporal Sampling fMRI Study.” *European Journal of Neuroscience* 20, no. 9: 2447–2456. <https://doi.org/10.1111/j.1460-9568.2004.03687.x>.
- Zangl, R., and D. L. Mills. 2007. “Increased Brain Activity to Infant-Directed Speech in 6- and 13-Month-Old Infants.” *Infancy* 11, no. 1: 31–62. https://doi.org/10.1207/s15327078in1101_2.
- Zatorre, R. J. 2022. “Hemispheric Asymmetries for Music and Speech: Spectrotemporal Modulations and Top-Down Influences.” *Frontiers in Neuroscience* 16: 511. <https://doi.org/10.3389/fnins.2022.1075511>.
- Zhang, J., Y. Meng, C. Wu, Y. T. Xiang, and Z. Yuan. 2019. “Non-Speech and Speech Pitch Perception Among Cantonese-Speaking Children With Autism Spectrum Disorder: An ERP Study.” *Neuroscience Letters* 703: 205–212. <https://doi.org/10.1016/j.neulet.2019.03.021>.
- Zhang, Y., T. Koerner, S. Miller, et al. 2011. “Neural Coding of Formant-Exaggerated Speech in the Infant Brain.” *Developmental Science* 14, no. 3: 566–581. <https://doi.org/10.1111/j.1467-7687.2010.01004.x>.
- Zhou, X., L. Wang, X. Hong, and P. C. M. Wong. 2024. “Infant-Directed Speech Facilitates Word Learning Through Attentional Mechanisms: An fNIRS Study of Toddlers.” *Developmental Science* 27, no. 1: e13424. <https://doi.org/10.1111/desc.13424>.
- Zwaigenbaum, L., S. Bryson, T. Rogers, W. Roberts, J. Brian, and P. Szatmari. 2005. “Behavioral Manifestations of Autism in the First Year of Life.” *International Journal of Developmental Neuroscience* 23, no. 2–3: 143–152. <https://doi.org/10.1016/j.ijdevneu.2004.05.001>.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.