CHARACTERISTICS OF VOCAL INTERACTIONS BETWEEN PARENTS AND CHILDREN WITH AUTISM SPECTRUM DISORDER IN RELATION TO HIERARCHICAL TEMPORAL CLUSTERING

by

Yumeng Zhang

Advised by Miriam Lense, Ph.D.

Presented to the Department of Medicine, Health, and Society

at Vanderbilt University

in Partial Fulfillment of the Requirements

for the Degree of Bachelor of Arts with Honors

Vanderbilt University

Nashville, TN

April 2021

ABSTRACT

The current study explores the vocal interactions of parents and children with and without Autism Spectrum Disorder (ASD) through a novel measure of hierarchical acoustic clustering in order to better understand the specific characteristics of parent and child speech that differ within ASD communication. Interactions between the parent and child were video-recorded and coded for the frequency and timing of parent and child vocalizations. To measure hierarchical temporal clustering, audio recordings of the dyadic interactions were analyzed across twelve timescales ranging approximately from the phoneme-level scale to the phrase-level scale and quantified using Allan Factor (AF) variances. There were three main findings of the study. First, significant relationships were found between frequency of interpersonal turns and ASD toddler language and developmental assessment scores. Second, ASD dyads exhibited significantly greater hierarchical temporal clustering compared to TD dyads. Third, the vocal characteristics which most correlated with hierarchical clustering in ASD dyads were frequency of total vocalizations and turn-taking, particularly when considering total interpersonal turns and total turns. These findings call attention to the importance of turn-taking in communication, and the reduced quality of turn-taking in interactions of parents and children with ASD.

ACKNOWLEDGEMENTS

I would like to thank my advisor, Dr. Miriam Lense, for all her support and mentorship throughout the whole process. I would also like to thank the lab members at the Music Cognition Lab, especially Olivia Boorom for her unparalleled guidance and Talia Liu for being such an integral part of this thesis. Thank you to the parents and toddlers that participated in this project, and lastly to my family and friends for their continued support.

TABLE OF CONTENTS

INTRODUCTION

Impairments in social communication and language are the most prominent characteristics of Autism Spectrum Disorder (ASD), in addition to restricted and repetitive patterns in behaviors, interests, and activities (American Psychiatric Association, 2013). Although core deficits in communication and social interaction appear towards the end of the first year of life (Baranek, 1999; Osterling, Dawson, & Munson, 2002; Osterling & Dawson, 1994; Ozonoff et al. 2010; Zwaigenbaum et al. 2005), the average age of ASD diagnosis remains around four years of age (CDC, 2020). Early presenting features of ASD include impairments in both verbal and nonverbal communication (American Psychiatric Association, 2013). Expressive and receptive language impairments are common features of ASD (Boucher, 2012; Gamliel, Yirmiya, Jaffe, Manor, & Sigman, 2009; Gernsbacher, Geye, & Ellis Weismer, 2006), however, the heterogeneity in the presentation and severity of ASD as well as the level of functioning in individuals with ASD must be noted. Currently, about 1 in 54 children in the United States meet the criteria for ASD (CDC, 2020). Because early intervention can alter the course of ASD, it is crucial to be aware of the development and early signs of ASD.

Literature Review

The current study will investigate the interactions of children with and without ASD through the lens of hierarchical temporal patterns of acoustic information. This novel measure of parent-child vocal interactions will shed light on the differences in communication between parents and children with and without ASD. Understanding these differences can be informative for interventions that utilize targeted communication strategies for parents of children with ASD. The following sections will provide an overview of the existing literature regarding differences in language between ASD and typically developing (TD) children, parent usage of infant-directed speech, the role of the parent in child language development, and hierarchical temporal clustering. This background is necessary to understanding the main focus of the current study, which aims to investigate the relationship between hierarchical temporal patterns and vocalization characteristics of interactions between parent and child.

a. Language Development in Children with ASD

Studying vocal development in children with ASD has the potential to provide clinical opportunities for enhancing language trajectories and outcomes. Studies have found a significantly lower rate of vocalization and speech like vocalizations in children with ASD compared to TD children (Warlaumont, Richards, Gilkerson, and Oller, 2014; Plumb & Wetherby, 2013). A few studies have shown that toddlers with ASD produce significantly more non-speechlike vocalizations than their TD peers (Plumb & Wetherby, 2013; Schoen, Paul, and Chawarska, 2011). Because parents are more likely to respond to speechlike rather than non-speechlike vocalizations, reduced speech-like vocalizations from children with ASD can impact the back and forth nature of parent-child communication.

In the study by Plumb and Wetherby (2013), vocalizations were coded from recorded behavior samples of 125 children between ages 18 and 24 months, divided into 50 toddlers with ASD, 25 with a developmental delay (DD), and 50 TD toddlers. The children who were later diagnosed with ASD produced a significantly lower percentage of speech-like vocalizations and a higher percentage of atypical vocalizations compared to the TD group. However, no differences were found between ASD and DD groups for these measures, and speech-like vocalizations between the ASD and TD groups did not differ in syllabic complexity.

Schoen, Paul, and Chawarska (2011) compared 30 toddlers with ASD, 11 typically developing age-matched (TDA) controls, and 23 typically developing language-matched (TDL) controls, between the ages 18 and 36 months. Results showed that while ASD toddlers produced the same total number of vocalizations as the other groups, they produced significantly more nonspeechlike vocalizations than the TDA group. Compared to the TDA group, significantly fewer words and word approximations were produced by the ASD group, however, both groups produced the same percentage of correct consonants.

A study by Chenausky, Nelson and Tager-Flusberg (2017) included 18 toddlers at low-risk (LRC) for ASD, 18 high-risk siblings without ASD (HRA-), and 10 high-risk siblings with ASD (HRA+). Vocalization rate and number of different consonants were obtained from 30-minute speech samples. The results found that HRA+ toddlers consistently produced a lower vocalization rate, but no differences in non-speech-like vocalization rate. HRA-, rather than HRA+ toddlers, had the least number of different consonants. Overall, lower vocalization rate was not associated with reduced number of consonants in this study.

As research continues to expand in this field of language development in children with ASD, conflicting evidence arises as well. While studies have found that toddlers with ASD produce fewer speech-like vocalizations, it is unclear if non-speech-like vocalization rate is higher than TD peers. Further studies must investigate the specific nature of speech-like vocalizations, and hone in on variables predictive of vocalization rate in young children. The complexity of early vocalizations can be investigated as an avenue to understand the differences in ASD vs TD language development.

A measure of vocal complexity can be indicated by the number of different consonants produced. Consonant production delays have been found in children with ASD, however conflicting evidence exists. Early studies of children with delays in expressive language have determined a relationship between vocalization rate and phonetic delay (Paul & Jennings, 1992; Rescorla & Ratner, 1996). These studies have led researchers to suggest that consonant production delays may be the result of less vocal practice, due to the reduced opportunities for articulatory practice and auditory feedback from a communication partner (Pharr, Ratner, & Rescorla, 2000). A study by Talbott (2015) found that the rate of consonant-vowel production at 9 months correlates with expressive language at 12 months in infants with ASD. These findings provide evidence that early vocal production predicts later language abilities in children with ASD, in addition to TD children. In a more recent study, Chenausky, Nelson & Tager-Flusberg (2017) found that children with ASD did not produce a significantly lower number of consonants compared to TD and high-risk groups, however, language trajectories measured by standardized language assessments of expressive and receptive language differed between all three groups.

In addition to consonant production, canonical babbling can also be studied as a marker of early vocal production. Canonical babbling, characterized by repeated syllables of consonant and vowels, usually appears around 5-10 months of age in typically developing children (Oller, 2000; Eilers and Oller, 1994). In infants with ASD, late babbling is a common characteristic. The absence of canonical babbling and persistence of precanonical babbling beyond 10-12 months may indicate abnormal vocal development and future speech and language problems (Oller et al. 1998; Lohmander et al. 2017). A longitudinal study reported that infants who did not produce canonical syllables at 10-12 months also had a reduced expressive vocabulary at 18, 24, and 30 months of age (Oller et al. 1999).

These measures of vocal development— vocalization rate and consonant production— are important for the study of early infant language development. Because vocal development is a precursor to language, it is useful to determine if these early measures of speech production are predictive of language delays in ASD. Improving vocalization complexity could potentially facilitate language development in ASD; increases in complexity can signal greater attunement to language learning and be indicative of attempts to say words infants understand but cannot accurately produce (McDaniel et al., 2020a). Increasing the rate of vocalization or vocal complexity might elicit more frequent adult responses to scaffold the child's language learning (McDaniel et al., 2020a). A recent study of the validity of vocal variables including vocal communication and complexity supports the use of these variables as predictors of expressive language in children with ASD (McDaniel et al. 2020a, 2020b). As a response to early language development, parents often alter how they communicate with their infants, called infant-directed speech.

b. Infant-Directed Speech

When speaking to their infants, parents alter visual, tactile, and auditory cues in order to capture their infant's attention (Bruner, 1983; Stern, 1985). Infant-directed speech (IDS) is a speaking style often used when interacting with young infants. In contrast to adult-directed speech, IDS is characterized by acoustic features including heightened pitch, exaggerated pitch contours, slow tempo, and increased rhythmicity (Ferguson, 1964; Fernald, 1991; Katz, Cohn, & Moore, 1996; Papousˇek, 1992; Stern, Spieker, & MacKain, 1982; Trehub, Trainor, & Unyk, 1993). These vocal cues of IDS are very important for infants' socioemotional, behavioral, cognitive, and linguistic development (Spinelli, Fasolo, & Mesman, 2017). Infants are highly responsive to the features of infant-directed speech, more so than adult-directed speech (Cooper & Aslin, 1990; Kaplan, Goldstein, Huckeby, Owren, & Cooper, 1995; Pegg, Werker, & McLeod, 1992; Werker, Pegg, & McLeod, 1994). Functions of IDS include facilitating language acquisition, communicating affect, promoting social interaction, and attracting infants' attention (Golinkoff, Can, Soderstrom, & Hirsh-Pasek, 2015).

Evidence shows that IDS has the potential to enhance infants' language learning (Bernstein Ratner, 1986; Fisher & Tokura, 1996; Kemler Nelson, Hirsh-Pasek, Jusczyk, & Cassidy, 1989; Kuhl et al., 1997). It has been hypothesized that the exaggerated intonation and rhythmic properties of IDS may engage infants' attention to speech more efficiently than adult-directed speech (ADS) (Fernald et al., 1989; Garnica, 1977; Thiessen, Hill, & Saffran, 2005). In a recently conducted study, investigators honed in on the specific acoustic properties of IDS that connect to attention processing and perceptual learning— finding that the variable prosodic contours of IDS contain greater prosodic information than ADS, enhancing language learning in infancy (Räsänen et al., 2018). Other studies demonstrate that the exaggerated pitch contours of IDS facilitate vowel discrimination during the developmental stage when infants are learning vowel categories (Trainor & Desjardins, 2002). Additional evidence suggests that IDS boosts infant language learning by accelerating vocabulary growth and enhancing speech processing (Saffran, Aslin, & Newport, 1996; Weisleder & Fernald, 2013).

In addition to the role of infant-directed speech in infant language learning, studies have investigated infant preference and parent usage of IDS. Evidence suggests that infants with ASD attend to IDS less than TD infants (Franchini et al., 2017). Along with those findings, other studies have determined that the amount of infant attention to IDS at 12 months predicts later expressive vocabulary (Vouloumanos & Curtin, 2014). Additional studies have found that parents of infants with ASD change their behavioral patterns of communication and use more IDS compared to parents of TD infants (Cohen et al., 2013). These studies suggest that infant-directed speech may be an important scaffold to enhance early vocalization development in infants with and without ASD.

c. Role of Parental Input in Infant Language Development

The interactions between children and their parents are crucial to infant development and socialization. Learning to produce speech-like vocalizations is an essential component of language learning and development (Oller, 2000). According to the social feedback loop, infant speech-language development is shaped by caregiver responses (Warlaumont et al., 2014). This hypothesis proposes that when a child produces a speech-like sound, they are more likely to receive an immediate, positive response from the caregiver than if the child vocalization was not speech-related. In turn, receiving an immediate response from the parent encourages the production of more speech-related vocalizations. The social feedback loop is supported by research suggesting that when the caregiver responses are vocal, their children's future vocalizations acquire acoustic characteristics resembling adults, including more speech-like properties, more vowel resonance, or better consonant-vowel timing (Bloom, 1988; Goldstein & Schwade, 2008). Individual interactions build upon each other over time, scaffolding and enhancing infant language development. Evidence suggests that a mother's responsiveness to her child's communicative behaviors predicts later language performance for both typically developing (Tamis-Lemonda, Bornstein, & Baumwell, 2001) and non-typically developing children (Yoder & Warren, 1999; Girolametto, 1988). Additionally, many studies support that greater vocal coordination between infant and parent is predictive of later language, cognitive, and perceptual abilities (Jaffe et al. 2001; Greenwood et al. 2010).

While the social feedback loop can be helpful in understanding typical development, it can also provide a novel perspective on language development in children with ASD. ASD is characterized by both atypical language development and social interaction. Children with ASD tend to produce fewer speech-related vocalizations (Warren et al., 2010; Paul et al., 2011; Patten et al., 2014), their vocalizations tend to be atypical (Sheinkopf et al., 2000; Oller et al., 2010; Peppé et al., 2007; Sheinkopf et al., 2012; Patten et al., 2014), and they acquire language at a slower rate compared to same age peers (Anderson et al., 2007). ASD can impact the social feedback loop in three different ways. First, fewer speech-like vocalizations provide infants with ASD less opportunities for adult feedback (Warlaumont et al., 2014). Second, parent responses to an infant with ASD might be less contingent on the infant vocalizations, possibly due to the reduced social quality of the vocalizations (Warlaumont et al., 2014). Third, social impairment can hinder the ability of children with ASD to learn from caregiver feedback (Warlaumont et al., 2014).

d. Temporal Acoustic Structure of Parent Speech

Temporal attention to prosodic information may scaffold language acquisition in infants (Diego-Balaguer et al. 2016). Parents modulate the temporal structure of utterances when speaking or singing with infants compared to adults (Falk & Kello, 2017). Acoustic energy in speech signals is arranged hierarchically and can be expressed as clusters of varying duration nested over a range of timescales (Abney, Paxton, Dale, & Kello, 2014; Luque, Luque, & Lacasa, 2015). In shorter timescales, phonemic variations (20-100ms) are nested within syllables and words (100-500ms), which are nested within phrases (500-4000ms), and which are nested within utterances (1000-6000ms) (Falk & Kello, 2017). Clusters emerge when analyzing patterns of temporal events, defined as discrete points in time that designate a significant modulation of acoustic energy (Falk & Kello, 2017). Allan Factor (AF) Analysis is a new measure used to quantify the variability of temporal events across multiple timescales in conversational speech (Allan, 1996; Abney, 2014; Falk & Kello, 2017). Abney et al. (2014) reported that AF analyses of temporal event clustering are sensitive to different speech styles— argumentative conversations showed greater temporal clustering and structure relative to affiliative ones, at longer timescales.

In recent studies comparing IDS and ADS using AF analyses, evidence shows enhanced hierarchical structure in infant-directed speech and song, indicating that nested clustering of temporal events corresponds with increasing temporal variability at hierarchical levels of linguistic structure (Falk & Kello, 2017). These results help reveal clear differences between the temporal structures of ID and AD speech and singing. The cause of these differences is hypothesized to be attributed to certain characteristics of speech. Several studies have found that boundary positions of speech are more likely to attract infants and assist with segmenting and remembering words (Trainor & Adams, 2000; Seidl & Johnson 2006; Shukla, White, & Aslin, 2011). Features of ID speech tend to become heavily boundary-oriented (Wang et al., 2016) and contain more important words and concepts before boundaries, most likely to reinforce infants' memory and learning of words (Fernald, 2000; Fernald & Mazzie, 1991). There is also evidence that adults adapt the amount of nested clustering in their speech to match the clustering in infant vocalizations (Abney et al. 2016). This evidence for coordination in infant-parent dyads suggests that adults adapt the complexity of language directed to their child in response to the properties of their child's language (Abney et al. 2016; Snow, 1989; 1995).

While this work was groundbreaking in discovering these novel acoustic features of IDS, there is still much to be uncovered, particularly in clinical populations. The current study will use this new metric, Allan Factor, to analyze the differences between parent speech targeted to typically developing children and children with ADS. This study will also look at how temporal acoustic structure relates to children's language levels, as well as vocal characteristics of parent speech.

The Current Study

This study expands upon work previously done on the coordination of parent speech to typically developing populations or infants with ASD. While previous research has revealed the various ways parents adjust their language when interacting with their children, no study has investigated the relationship between the hierarchical temporal structure of parent-child communication and behavioral data from interactions between parents and children with ASD. Thus, this research aims to investigate temporal acoustic patterns in parent interactions with children with ASD in relation to measures of child language complexity and the acoustic characteristics of vocalizations between parent and child. These findings will uncover how parents and children adjust characteristics of their vocal interactions depending on the developmental level of the child.

The current study has two aims. The first aim will compare temporal clustering of speech in parent-child dyads of ASD versus TD toddlers. The second aim will focus on the ASD dyads and examine behavioral measures of vocalization amount and coordination— vocalization frequency, duration, and turn-taking— to determine which characteristics impact temporal clustering over various timescales. Based on previous work, I hypothesize that TD and ASD parent-child dyads will exhibit significant differences in hierarchical temporal clustering patterns. In comparison to TD dyads, ASD dyads will exhibit greater hierarchical clustering and greater variability and duration in interpersonal turns, as parent speech will more greatly resemble infant directed speech.

METHODS

Participants

Participants of this study included 21 TD parent-child dyads and 15 ASD dyads. At the time of the study, the 21 children of the TD sample were between 18 and 21 months (M = 18.57 months, SD = .98 months). The language characteristics of the TD sample are listed below in Table 1. The 15 children of the ASD sample were between 18 and 34 months (M = 27.47 months, SD = 4.75 months). All participants indicated English as the primary language at home. Parent education ranged from 9th-11th grade to a professional degree (MD, PhD, JD), with courses towards college as the median level. Participant characteristics, including ADOS-2 comparison scores and Mullen Scales of Early Learning age equivalences are provided in Table 2.

These participants were drawn from an ongoing study investigating parent-child interactions conducted at Vanderbilt University Medical Center. Interested parents for the larger project were contacted and screened for eligibility. As part of the larger study, toddlers were seen in the laboratory over several visits where they participated in a range of standardized behavioral assessments and parent-child play sessions. The present study focuses on toddlers who received a diagnosis of ASD during the assessment portion of the study. Autism diagnosis for all participants was established via clinical best estimate through a comprehensive diagnostic evaluation, including assessment of developmental functioning (Mullen Scales of Early Learning (MSEL; Mullen, 1995)) and the Autism Diagnostic Observation Schedule, Second Edition (ADOS-2), (Lord et al., 2012). Written informed consent was obtained by all parents, and compensation was provided for their participation in the study. Ethical approval was provided by the institutional review board of the university.

Table 1. ASD and TD Sample Language Characteristics

Sample	ASD		TD	
	M	*SD*	*M*	*SD*
Chronological Age (months)	27.47	4.75	18.57	0.98
MSEL Visual Reception	20.87	4.03	21.20	4.14
MSEL Expressive Language	13.93	5.82	19.52	1.86
MSEL Receptive Language	11.87	6.27	21.55	4.11
Language Composite	13.97	5.53	20.60	2.62

Note: Language composite = Average of MSEL Expressive and Receptive Language

Table 2. ASD Toddler Demographics

Participant	*Sex*	*Age (months)*	*Race*	*ADOS-2 CS*	*MSEL EL*	*MSEL RL*	*MSEL VR*	*CDI Words Produced*	*CDI Words Understood*
1	**M**	**18**	**Mixed**	**5**	**17**	**14**	**19**	**29**	**251**
2	**M**	**31**	**White**	**10***	**8**	**6**	**18**	**34**	**94**
3	**M**	**22**	**White**	**10**	**8**	**4**	**13**	**0**	**82**
4	**M**	**34**	**Mixed**	**8***	**23**	**30**	**27**	**149**	**203**
5	**M**	**27**	**Black**	**10**	**10**	**6**	**20**	**16**	**90**
6	**M**	**25**	**Mixed**	**8**	**12**	**9**	**20**	**5**	**119**
7	**M**	**32**	**White**	**9***	**24**	**13**	**26**	**71**	**157**
8	**M**	**24**	**White**	**10**	**15**	**11**	**21**	**6**	**73**
9	**F**	**30**	**White**	**8**	**16**	**9**	**27**	**69**	**101**
10	**M**	**32**	**White**	**7***	**20**	**19**	**24**	**20**	**70**
11	**M**	**24**	**White**	**10**	**7**	**10**	**16**	**0**	**49**
12	**M**	**32**	**White**	**7***	**7**	**14**	**19**	**0**	**5**
13	**M**	**26**	**White**	**10**	**8**	**10**	**20**	**0**	**4**
14	**M**	**32**	**White**	**9***	**17**	**13**	**24**	**100**	**127**
15	**M**	**23**	**White**	**9**	**17**	**10**	**19**	**23**	**192**

Mean	**27.47**	**8.67**	**13.93**	**11.87**	**20.87**	**34.80**	**107.80**
SD	**4.75**	**1.50**	**5.82**	**6.27**	**4.03**	**43.98**	**69.81**

*Note: ADOS-2 CS = ADOS-2 Comparison Score. Comparison Scores range from 1 to 10 with 10 being the highest level of autism-related symptoms. *The ADOS-1 was administered for participants 2, 4, 7, 10, 12, and 14. MSEL EL, RL, VR = MSEL Expressive Language, Receptive Language, Visual Reception. MSEL scores are age equivalences. CDI scores are raw scores.*

Measures

Autism Diagnostic Observation Schedule

The Autism Diagnostic Observation Schedule, Second Edition (ADOS-2; Lord et al. 2012) is administered to toddler participants to assess and diagnose autism spectrum disorders. The ADOS-2 is a standardized assessment of social, communication, and imagination skills across all ages and developmental levels. A qualified and trained examiner uses the ADOS-2 to measure children's language, social communication, and restricted or repetitive behaviors during a semi-structured play session with the child. Empirically derived cutoffs determine if the child meets the criteria for ASD and reliably distinguishes between autism, autism spectrum disorders, other non-ASD disorders, and typically developing children.

MacArthur-Bates Communicative Development Inventory

Prior to the play session, parents completed the Words and Gestures form of the MacArthur-Bates Communicative Development Inventory (CDI; Fenson et al. 2007). The CDI is a commonly used measure of expressive and receptive vocabulary and grammar. The Words and Gestures form consists of two parts. Part I is a 680-word vocabulary checklist organized into nineteen semantic categories that asks parents to indicate the words as understood and/or used by their child. In Part II, parents are asked to record communicative and symbolic gestures attempted or completed by their child. The form is scored and organized into percentiles based on sex. For the present study, raw scores for Words Produced and Words Understood were utilized from these assessments.

Mullen Scales of Early Learning

Toddlers were also administered the Mullen Scales of Early Learning (MSEL; Mullen, 1995). The MSEL is a standardized developmental assessment of cognition, language, and motor functioning. It is organized into five subscales: (a) gross motor, (b) fine motor, (c) visual reception (non-verbal problem solving), (d) receptive language, and (e) expressive language. Age equivalence scores and T-scores are provided for the subscales. The current study utilizes the visual receptive and expressive and receptive language age-equivalence scores.

Procedure

Free Play

Parents and toddlers were video-recorded during a free play session during the study visit. During the free play session, the child and parent were seated on the floor in the center of the room and asked to talk and play together. Various quiet toys and books were provided to the dyad to play with, including books, balls, and animal figures and a toy barn. For a full list of available toys, see Appendix A. Four cameras used to videotape the session were placed at various locations to optimally capture the parent and child during play. Parents were asked before the session to sit facing the cameras if possible. To capture audio, an AKG Tascam DR-40 microphone was placed on a table facing the dyad.

Coding

Data was derived from a 7-minute section of the recorded parent-child free play sessions. The first two minutes were not included in the data analysis in order to give the parent and child time to settle into play and become comfortable in the new setting, and the last few minutes of the interaction were excluded due to a few toddlers who became upset towards the end. The first pass of coding was done using a time-locked annotation program (ELAN; Brugman and Russel, 2004) that allows for segmentation and annotation of the data from a video and audio source. Both parent and child vocalizations were isolated based on phrase and breath group (Oller, 1973). When annotating the file, the vocalizations were identified as linguistic or non-linguistic. Toddler linguistic vocalizations were coded as “IL” and parent linguistic vocalizations were coded as “PL”. A linguistic vocalization is defined as a non-vegetative voiced sound created by vibrating vocal folds. Non-linguistic vocalizations included (a) voiced laughing, sighing, or crying, (b) reflexive vegetative sounds such as burps, hiccups, or coughs, (c) whispers, or (d) raspberries, clucks, or kisses. Rhythmic toy noises, including banging, ball dropping, and clapping, were identified and annotated.

Reliability

The primary and secondary coders were trained until they reached 85% reliability on at least three consecutive videos. To assess inter-coder reliability, the secondary coder independently coded a random sample of 20% of the sessions (N = 3).

Data Analysis

Hierarchical Temporal Clustering

The first aim of this study was to compare hierarchical temporal clustering between ASD and TD dyads. The audio recordings of the dyadic interactions were analyzed across 12 timescales ranging approximately from the phoneme-level scale to the phrase-level scale and quantified using Allan Factor (AF) variances. Hierarchical clustering was measured using Allan Factor (AF) variances, which are coefficients of variance with respect to adjacent timescales (Falk & Kello, 2017). The degree of nested clustering was measured using the slope of the best fit regression line of the AF variances. AF variances were used to quantify dyadic interactions across twelve timescales. A quadratic equation was fit to each dyad's AF results across all timescales to quantify patterns of temporal clustering. Steeper slopes correspond to greater nesting. Only the linear component of the AF variance will be reported.

Vocalization Variables

The second aim of this study was to examine characteristics of vocal interactions associated with individual differences in dyadic temporal clustering. Conversational variables were extracted from the ASD dyad interactions, and include parent and infant vocalization frequency, average vocalization duration, total vocalization duration, and turn-taking variables including interpersonal turns and latencies between parent and infant linguistic vocalizations and all vocalizations. The list of variables and their definitions and calculations are located in Appendix B. In order to determine their relationship, AF slopes were correlated with vocalization variables using Pearson Correlation tests.

RESULTS

The current study was designed to investigate characteristics of vocal interactions between parent and child, specifically measures of vocalization amount and coordination – vocalization frequency, duration, and turn taking – associated with individual differences in dyadic hierarchical temporal clustering.

After preliminary analyses are presented to provide descriptive information about the vocal interactions, results will be presented in the order of the two study aims. First, results comparing hierarchical temporal clustering between TD and ASD dyads will be presented. This is followed by results exploring the behavioral measures of vocal coordination correlated with hierarchical temporal clustering in ASD dyads. All analyses were carried out using the R Studio (RStudio Team, 2020).

1. Preliminary Analyses

Information regarding the vocalization variables is provided in Table 3 and 4, including means and standard deviations. The preliminary information in Table 3 provides an overview of the differences between parent and infant vocalizations. For example, while both parents and infants are vocalizing during the interaction, parents are vocalizing more frequently and for longer durations. Thus, most of the acoustic signal captured during the interaction is from the parent. In Table 4, information about the dyad is organized into linguistic vocalizations and all vocalizations. All vocalizations include both linguistic and non-linguistic vocalizations, and was included in the analyses to capture a broader acoustic signal of the interaction. Linguistic vocalizations may better represent communicative behaviors across the dyads, however, the acoustic signal used for the Allan Factor analyses of acoustic temporal clustering does not discriminate between communicative or non-communicative vocalizations.

Preliminary analyses were conducted to determine the potential effects of ASD toddler characteristics on vocalization variables. Table 5 displays the correlations between conversational variables and toddler characteristics including age, MSEL Visual Reception (VR), Receptive Language (RL), Expressive Language (EL), Language Composite, and ADOS Comparison Score. Pearson correlation tests revealed significant relationships between frequency of infant-to-parent turns and MSEL visual reception ($r(13) = .53$, $p = .043$) and expressive language ($r(13) = .52$, $p = .046$), and language composite ($r(13) = .51$, $p = .05$). No other correlations were statistically significant. Thus, parents more frequently responded to vocalizations of children with higher developmental and language levels.

Table 3. Vocalization Variables

	M	*SD*
Infant Frequency	68.6	39.84
Parent Frequency	149	38.06
Infant Duration	1.09	0.19
Parent Duration	1.18	0.22
Infant Total Duration	76.55	0.22
Parent Total Duration	173.98	48.29
Infant to Parent Turns	23.8	13.14
Infant to Parent Latency	0.69	0.19
Parent to Infant Turns	25.2	0.77
Parent to Infant Latency	0.77	0.15
Infant to Infant Turns	12.6	12.11
Infant to Infant Latency	0.9	0.52
Parent to Parent Turns	69.8	28.3
Parent to Parent Latency	1.1	0.14

Table 4. Vocalization Variables Across the Dyad

	All Vocalizations		Linguistic Vocalizations	
	M	*SD*	*M*	*SD*
Total Frequency	217.6	60.41	179.27	47.45
Total Duration	250.53	69.51	205.69	57.62
Total Interpersonal Turns	49	26.89	32.87	23.28
Total Intrapersonal Turns	82.4	24.49	66.07	24.41
Total Turns	131.4	35.08	98.93	30.33

Table 5. Correlations Among ASD Toddlers' Characteristics and Vocalization Variables

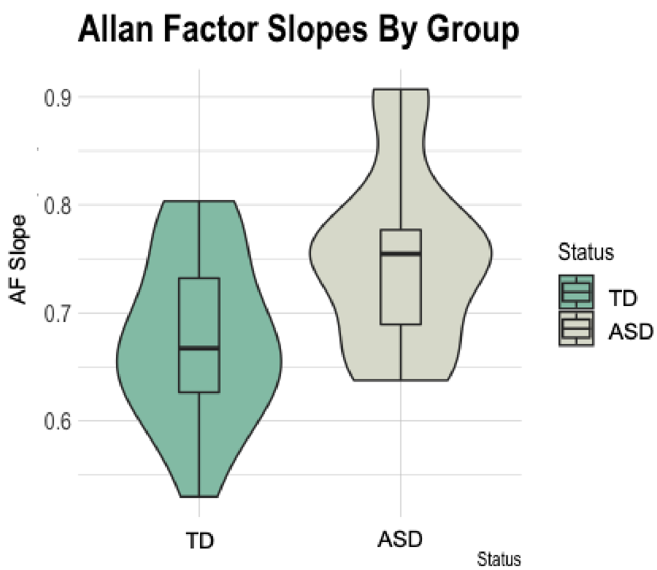
	Infant to Parent Turns	**Infant to Parent Latency**	**Parent to Infant Turns**	**Parent to Infant Latency**	**Infant to Infant Turns**	**Infant to Infant Latency**	**Parent to Parent Turns**	**Parent to Parent Latency**
Age	.4	-.23	.31	.16	.21	-.13	-.07	.07
VR	**.53***	-.27	.39	-.22	.02	.14	-.05	-.12
RL	.42	.16	.37	.29	-.01	-.14	-.06	.18
EL	**.52***	-.02	.47	-.1	.21	.01	-.16	-.24
Language Composite	**.51***	.08	.46	.11	.1	-.07	-.16	.02
ADOS-2	-.11	-.28	-.11	-.27	.23	.15	-.15	-.04

Note. N = 15. *** *p < 0.05. MSEL EL, RL, VR = Mullen Scales of Early Learning Expressive Language, Receptive Language, Visual Reception. All MSEL scores are age equivalences in months. ADOS-2 = ADOS-2 Comparison Score.*

2. Hierarchical Acoustic Temporal Clustering Between ASD and TD Dyads

A Welch two-sample t-test was conducted to compare the slopes of TD dyads to slopes of ASD dyads. Results indicate a significant difference in the linear component of the Allan Factor slope of the TD sample (M = .68, SD = .08) and the ASD sample (M = .75, SD = .08) (t(29) = -2.73, p = .01). As expected, the Allan Factor linear component was significantly steeper for the ASD dyads compared to the TD dyads. Therefore, ASD dyads demonstrated greater nested clustering of events compared to TD dyads.

Figure 1. TD vs ASD Toddler Allan Factor Slopes



3. Correlation of Dyadic Hierarchical Temporal Clustering with Characteristics of Vocal Interactions

The second aim of this study was to examine behavioral measures of vocal interactions associated with individual differences in dyadic temporal clustering. To examine correlations with AF slope, only frequency metrics from the vocal interactions were considered to be consistent with the Allan Factor measurement of event frequencies over various time scales. Correlations between AF slope and parent and infant vocalization variables are displayed in

Table 6, while correlations of total vocalization variables *across the dyad* are displayed in Table 7. The following results will expand upon vocalization frequency and turn-taking variables.

Table 6. Correlations Among Allan Factor Slope and Conversational Variables

Variables	**Allan Factor Slope**
Parent Frequency	**-.52***
Infant Frequency	-.41
Parent to Infant Turns	-.51
Infant to Parent Turns	**-.53***
Parent to Parent Turns	-.36
Infant to Infant Turns	-.25

Note. N = 15. * *$p < 0.05$*

Table 7. Correlations Among Allan Factor Slope and Dyadic Vocalization Variables

	All Vocalizations	Linguistic Vocalizations
	AF Slope	
Total Frequency	**-.52***	**-.64***
Total Frequency No SS	-.15	-.33
Interpersonal Turns	**-.53***	-.47
Intrapersonal Turns	**-.54***	-.41
Total Turns	**-.78****	**-.68****

Note. N = 15. * *$p < 0.05$, ** $p < 0.01$*

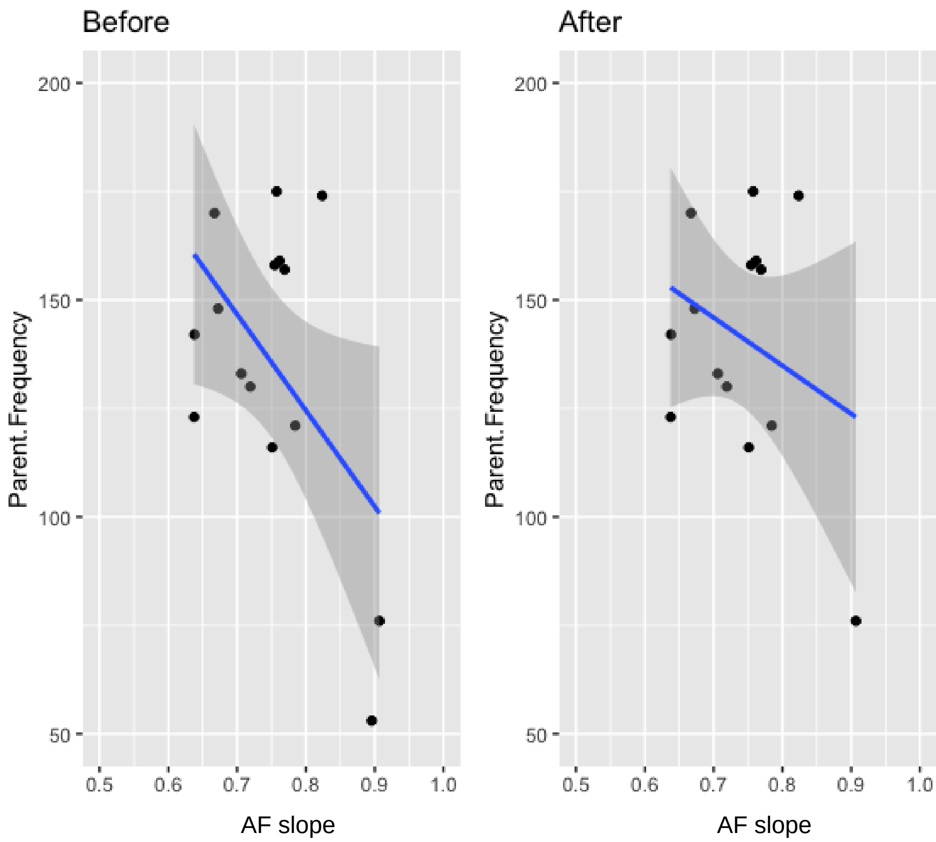
a. Vocalization Frequency

Infant, parent, and total vocalization frequency were correlated with the AF slope using Pearson correlations. Correlation test results revealed a significant strong negative correlation between AF slope and total vocalization frequency ($r(13) = -.52$, $p = .046$) and total linguistic vocalization frequency ($r(13) = -.69$, $p = .01$). Vocalization frequency across the dyad was strongly associated with the AF slope.

The following results focus on the relationship between AF slope and linguistic parent and infant frequency. Results revealed a significant negative correlation between parent vocalization frequency and AF slope ($r(13) = -.52$, $p < .05$) and a moderate negative correlation between infant vocalization frequency and AF slope ($r(13) = -.41$, $p = 0.13$).

Regression analysis was conducted on all vocalization frequency variables to determine the leverage and influence of potential outlier points (see Appendix C). In the relationship between parent frequency and AF slope, participant 2 displayed a significantly high influence point with a Cook's D value of .78. When removed from the data set, there was no significant effect of AF slope on Parent Frequency $b = -.30$, $t(13) = -1.09$, $p = .30$). Overall, while steeper AF slopes were initially associated with fewer parent vocalizations, this effect was driven by a specific dyad characterized by a very low number of parent vocalizations.

Figure 2. Correlation of Parent Frequency and AF Slope Before and After Removing Outlier



b. Turn-Taking

Interpersonal Turns

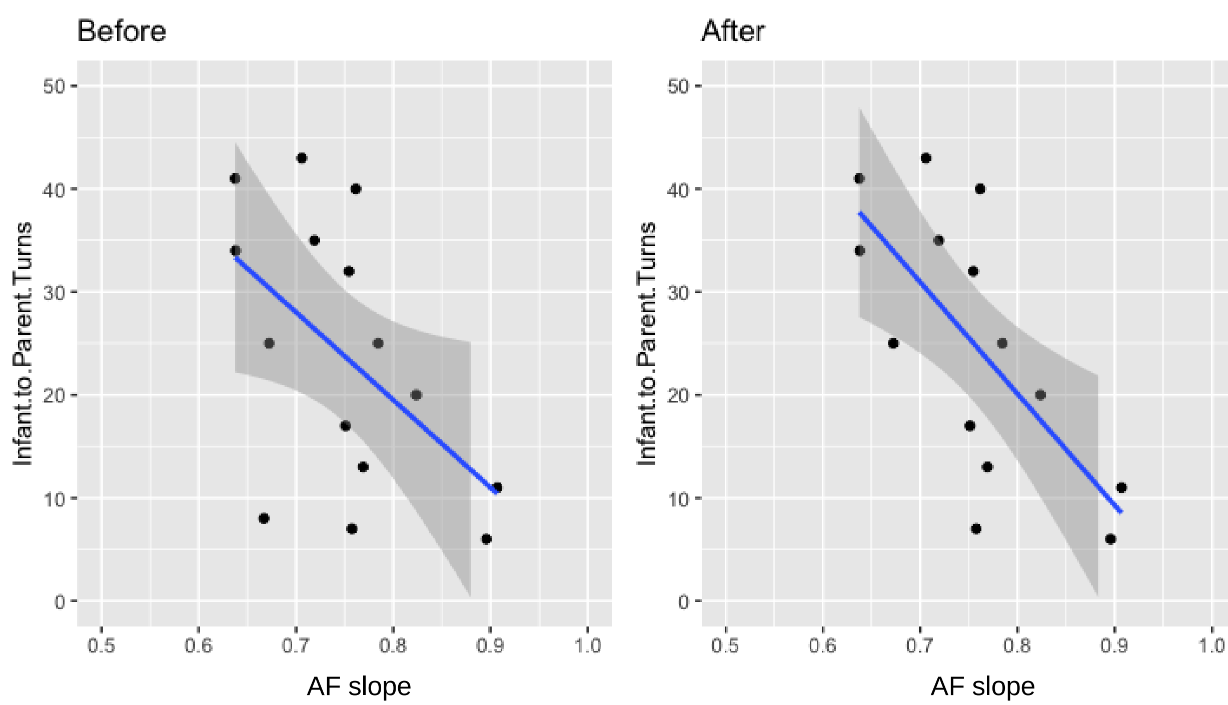
An interpersonal turn is the switch in conversation from one speaker to the other speaker, marked by a period of silence between speakers less than 3000ms long.

Total interpersonal turns across the dyad were calculated by the addition of parent-to-infant and infant-to-parent turns, for both linguistic and non-linguistic turns. Results of the Pearson correlation tests showed a significant strong negative correlation between total interpersonal turns and AF slope ($r(13) = -.53$, $p = 0.04$). After isolating linguistic interpersonal turns, a negative correlation that followed the same trend but did not reach conventional significance was found with total linguistic interpersonal turns and AF slope (r(13) = -.47, p = .079).

The following results focus on interpersonal turns done by either parent or infant. Pearson correlation tests showed a significant negative correlation between AF slope and linguistic infant-to-parent turns (r(13) = -.53, p = .045), and linguistic parent-to-infant turns (r(13) = -.51, p = .055).

Regression analysis was conducted on all interpersonal turns to determine the effect of influence and leverage (see Appendix C). In the relationship between AF slope and infant-to-parent turns, participant 12 displayed a significantly high influence point with a Cook's D value of .37. When removed from the data set, there was an even larger significant effect of AF Slope on infant-to-parent turns (b =-.69, $t(13) = -3.29$, $p = .006$). Therefore, higher AF slopes were observed for dyads with lower frequency of turn-taking.

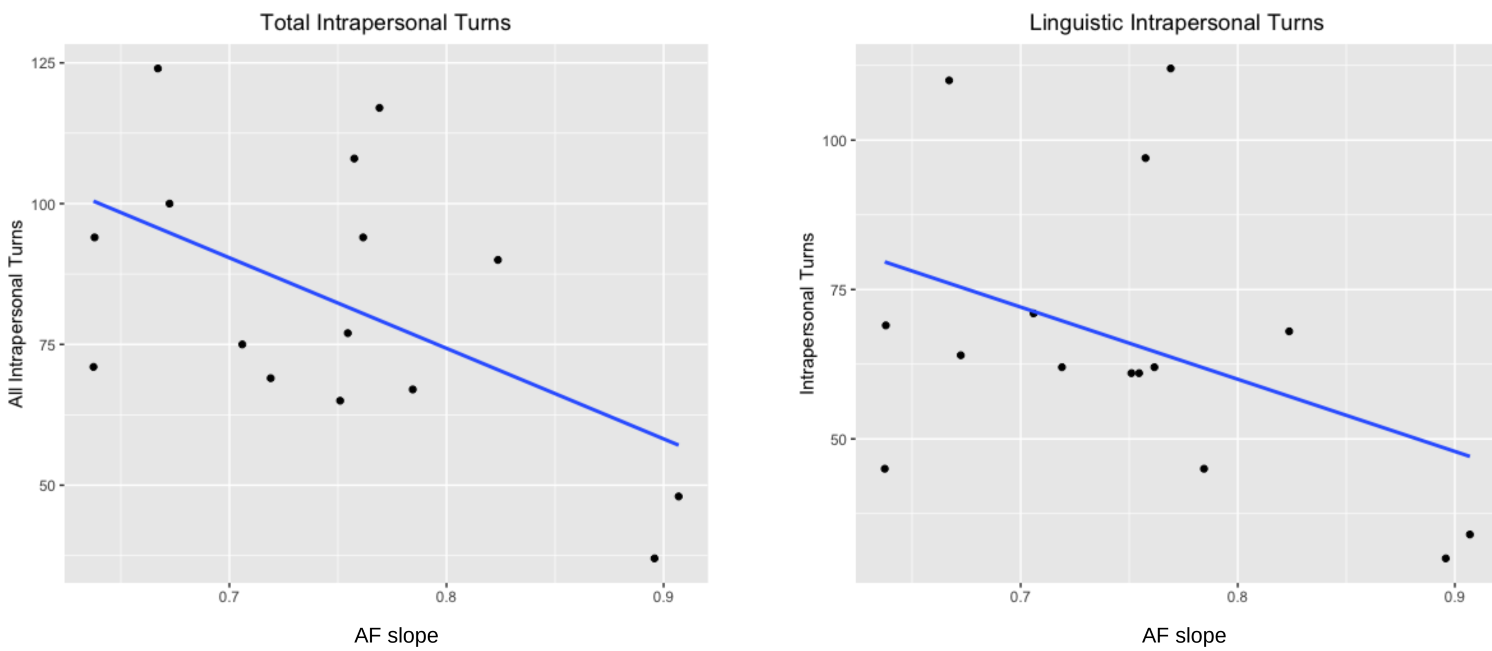
Figure 3. Correlation of Infant-to-Parent Turns and AF Slope Before and After Removing Outlier



Intrapersonal Turns

An intrapersonal turn is defined as a cycle of speaker vocalization, pause, same speaker vocalization. Total intrapersonal turns including all parent-to-parent and infant-to-infant turns did show significant negative correlations with AF Slope ($r(13) = -.54$, $p = .038$). A similar pattern was seen for linguistic intrapersonal turns with AF Slope though this did not reach conventional significance levels ($r(13) = -.41$, $p = .13$). No significant correlations existed for AF Slope and infant-to-infant or parent-to-parent turns, whether including all or only linguistic turns.

Figure 4. Total and Linguistic Intrapersonal Turns Correlation with AF Slope



Total Turns

Total turns are measured by the addition of interpersonal and intrapersonal turns, including all parent-to-infant, infant-to-parent, parent-to-parent, and infant-to-infant turns. Extremely strong negative correlations existed between AF Slope and all turns ($r(13) = .78$, $p < 0.001$) and linguistic total turns ($r(13) = -.68$, $p < 0.01$).

DISCUSSION

The purpose of this research study was to examine how vocal characteristics of the interactions between a child with ASD and their parent are coordinated with individual differences in dyadic hierarchical temporal clustering. The study investigated the differences in hierarchical temporal clustering between TD dyads and ASD dyads, and sought to determine which characteristics of vocal interactions in the ASD dyads correlated with these differences. There were 3 main findings. First, significant relationships were found between frequency of interpersonal turns and ASD toddler language and developmental assessment scores. Second, ASD dyads exhibited significantly greater hierarchical temporal clustering compared to TD dyads. Third, the vocal characteristics which most correlated with hierarchical clustering in ASD dyads were frequency of total vocalizations and turn-taking, particularly when considering total interpersonal turns and infant-to-parent turns. These main findings will be discussed in turn.

1. Hierarchical Acoustic Temporal Clustering Between ASD and TD Dyads

Based on previous research on hierarchical temporal clustering (Falk & Kello, 2017), I predicted greater event clustering in ASD parent-infant dyads compared to TD dyads. Infant-directed speech has been found to show greater hierarchical temporal clustering compared with adult-directed speech across multiple timescales (Falk & Kello, 2017). I hypothesize that parents of children with ASD often use more infant-directed speech in order to engage and communicate with a child who may not be reciprocating in conversation. Results from this study support previous research, and extend these findings to capture the differences in hierarchical temporal structure between ASD and TD dyads. This study found that ASD dyads showed greater hierarchical temporal clustering compared to TD dyads, as measured by steeper slopes characterizing the AF variance of nested clustering.

This novel acoustic measure of hierarchical acoustic structure across timescales can shed light on the differences in communication between parents of TD children and parents of children with ASD. Previous studies have determined infant-directed speech to display characteristics including greater boundary-oriented speech, shorter utterances, reduced prosodic complexity, and more frequent pauses compared to adult-directed speech (e.g., Albin & Echols, 1996; Martin et al., 2016; McMurray et al., 2013; Trainor et al., 1997; Wang et al., 2015). The current study demonstrates greater hierarchical clustering in speech directed to and from toddlers with ASD, potentially indicating increased prosodic variation and/or increased variability in temporal structure of within person and between people vocalizations. This enhanced variability in speech between parent and child may function to increase attention to important cues in conversation and maintain arousal levels. In order to prevent infants from disengaging, prosodic variation can make stimuli more unpredictable and in turn, more interesting. Contrast in temporal patterns can help infants stay attuned to the interaction and possibly enhance the discriminability in hierarchical acoustic structure (de Diego-Balaguer, Martinez-Alvarez, & Pons, 2016; Delavenne, Gratier, & Devouche, 2013). These features of speech can be utilized by parents to engage children with ASD that struggle with maintaining attention levels and facilitating interactions.

While these findings reveal that hierarchical temporal clustering may represent parents' use of techniques to facilitate communication with infants, more information is needed about the specific characteristics of dyadic interactions that are driving these differences in hierarchical clustering. The next section focuses on the vocal characteristics of interactions between parents and toddlers with ASD and their correlations with AF variance values that measure nested clustering.

2. Correlation of Hierarchical Temporal Clustering and Vocal Characteristics of Dyadic Interactions

Hierarchical clustering, represented by AF variance slopes, was significantly correlated with the following vocalization variables: total vocalization frequency, infant-to-parent turns, total interpersonal turns, and total intrapersonal turns. The following section will review the significant relationships between AF slope and frequency of total vocalizations and turn taking measures.

Vocalization Frequency

Total vocalization frequency is a measure of the total number of linguistic utterances made by both parent and child during the interaction. Results revealed a significant negative correlation between total vocalization frequency and AF slope. This indicates that at steeper AF slopes, there is a lower frequency of vocalizations across both partners in the dyad. The frequency of vocalizations is likely directly related to the frequency of turn taking or vocal coordination (described more below). Hierarchical clustering is event-based, meaning it is measured by the frequency of events across various timescales. Thus, it is unsurprising that AF slope would be associated with event *frequency*. Additionally, it is also unsurprising that total vocalization frequency across the dyad is highly correlated with AF slope, while individual parent and infant frequency show no significant relationships. AF variances capture the vocal interaction across the dyad and cannot distinguish between different speakers. However, parents are still vocalizing more frequently than children, representing a larger portion of the acoustic signal of the interaction and thus driving the relationship.

Turn-Taking Measures

Three measures of turn-taking frequency significantly correlated with AF slope: infant-to-parent turns, total interpersonal turns, and total turns.

Infant-to-parent turns, while measured by the shift in interaction from child to parent, can also be indicative of parent responsiveness. After the child vocalizes, three possible options can occur. First, the child can pause and keep speaking – also known as an intrapersonal turn. Second, the parent can respond to the infant after a pause of less than 3000ms – defined as an infant-to-parent interpersonal turn. Third, neither the parent or child speaks and longer than 3000ms of silence passes before a vocalization. In the second case, it is the parent that drives the interaction forward by contingently responding to the child. Therefore, this measure can give insight into how often the parent is responding to their child. AF slope and infant-to-parent turns showed a significant strong negative correlation of -.69 after removing the influential outlier. Thus, fewer infant-to-parent turns are found at steeper slopes, also indicating lower parent responsiveness at steeper slopes.

Total interpersonal turns includes both infant-to-parent turns and parent-to-infant turns. Total interpersonal turns was significantly correlated with AF slope, showing a strong negative correlation of -.54 for all interpersonal turns and a correlation of -.53 for linguistic interpersonal turns. Following the same trend, fewer interpersonal turns are associated with steeper AF slopes and greater hierarchical clustering. These results are supported by Abney et al. 2014, a study that found fewer interpersonal turns at steeper slopes associated with argumentative as compared to affiliative conversations. Interpersonal turns reflect the coordination between parent and infant in back and forth interaction. A greater number of interpersonal turns between parent and infant may be evidence of cohesive interactional dynamics. At steeper slopes, fewer turns can be indicative of a breakdown in communication between partners.

Total intrapersonal turns also showed a significant negative correlation with AF Slope, with a correlation of -.54, while total linguistic intrapersonal turns was moderately correlated with AF Slope. As AF variance is a measure of event-based clustering, total intrapersonal turns including both linguistic and non-linguistic vocalizations may be more correlated with this measure of hierarchical clustering. Intrapersonal turns, compared to interpersonal turns, are not as representative of coordination between parent and infant. Intrapersonal turns are less indicative of back and forth interaction, as an intrapersonal turn is characterized by the continuation of the interaction by one speaker. The variability in the relationship of AF slope to intrapersonal turn measures leads to difficulties interpreting the results, when individual parent-to-parent turns and infant-to-infant turns are not significantly correlated with AF slope. Therefore, future studies with a larger sample size and data set are needed to make evidence-based conclusions.

3. The Role of Turn-Taking in Communication and Infant Language Development

Turn-Taking Measures and Language Skills

A few significant relationships existed between language assessment scores and infant-to-parent turns. Infant-to-parent turns showed significant positive correlations with MSEL visual reception and expressive language, and a language composite averaging receptive and expressive language. This indicates that a greater frequency of infant-to-parent turns is associated with greater non-verbal language skills (visual reception) and verbal language skills (expressive language), which is also supported by the positive correlation with the language composite.

Infant-to-parent turns measures both parent responsiveness and coordination to infant speech. These findings provide support that parent responsiveness and coordination to infant speech are related to infant language development. Studies have found that parent responsiveness to infant communicative behaviors can predict later language development in non-typically developing children (Yoder, 1999; Girolametto, 1988), while other studies have shown that greater vocal coordination is predicative of later language, perceptual, and cognitive abilities (Jaffe et al. 2001; Greenwood et al. 2010). Additionally, children at higher language and developmental levels can provide more speechlike vocalizations and clearer communicative signals for the parent to respond to.

While our results cannot determine the predictive value of turn-taking on infant language development, they still support the evidence that parent coordination and responsiveness are directly related to infant language skills.

Turn-Taking as a Measure of Coordination and Responsiveness

The coordination of turn-taking is foundational for an interaction to run smoothly. Existing literature, including the Interaction Engine hypothesis, suggests that turn-taking develops early in infancy and is grounded in the ability to anticipate and recognize the conversational partner's intentions (Hilbrink et al. 2015). The ability to coordinate turns is based on the predictability of the interaction and the sensitivity of each partner to the timing of turns (Hilbrink et al. 2015). In hierarchical temporal structure of ASD dyads, steeper slopes represent greater variability and less predictability in interaction. Interactions that are less predictable lead to difficulties in anticipating the communicative intentions of partner. Children with ASD have difficulties with attention to speech, resulting in less attunement to the speech pattern. In light of this, ASD toddlers may engage in less interpersonal turns with their interaction partner due to the variability and unpredictability of interactions with their parent.

The social feedback loop provides a framework for understanding the role of both infant and parent in communication and infant language development. It highlights the importance of caregiver responses in shaping infant language development. The social feedback loop can also reveal how communication differs with ASD, and explain why fewer vocalizations and interpersonal turns are seen in interactions between parents and children with ASD. Studies have shown that children with ASD tend to produce fewer speech-like vocalizations (Warren et al., 2010; Paul et al., 2011; Patten et al., 2014). Parents could be responding to fewer infant vocalizations, which results in fewer interpersonal turns between child and parent (Warlaumont et al., 2014). Another explanation would be that the lower quality of infant vocalizations might prevent parents from responding in a meaningful way (Warlaumont et al., 2014). Lastly, social impairment of children with ASD might hinder their ability to learn from and respond to parent feedback. This loop can scaffold language development in TD children, but for dyads of parents and children with ASD, the loop can break down in many ways to hinder coordination. Fewer infant vocalizations can lead to less parent responses, and less parental input can lead to even fewer infant responses. The cyclical nature of the social feedback loop can drive down the number of vocalizations and interpersonal turns. Thus, the social feedback loop can reveal why lower frequency of vocalizations and interpersonal turns are found at steeper slopes.

4. Implications

This research study filled in a gap in existing literature regarding the correlation of hierarchical temporal clustering to behavioral measures of communication between a parent and child with ASD. Not only did this study use a novel measure of hierarchical temporal structure, but the findings also provided insight into the specifics of interactions within the timescales of nested clustering. By hand-coding vocalization variables rather than using automated systems, we were able to extract accurate and detailed characteristics of the interactions between parent and child. The findings of the study indicated that differences in communication of ASD dyads compared to TD dyads can be attributed largely to turn-taking. The particularly strong correlations between turn-taking frequency and AF values indicates that reduced turn-taking often observed in ASD impacts the entire temporal organization across multiple timescales of the parent-child interaction. Turn-taking in interactions is associated with coordination and responsiveness between parent and child, and it is also predicative of infant language skills. Furthermore, the relationship between turn-taking and hierarchical temporal clustering provides support for the importance of parental coordination and responsiveness for child language development. These findings point to the importance of turn-taking in communication, which can be incorporated as a strategy in parent-training interventions. Lastly, these findings demonstrate the promise of using Allan Factor variances as a measure of dyadic interactions between parent and child, particularly how often individuals are vocalizing and taking turns. Therefore, future studies can utilize this automated measure as a representative picture of parent-child interactions rather than traditional time-consuming measures that require hand-coding.

5. Limitations and Future Directions

The limitations that existed for this research study must be considered. First, the study sample was limited in size and variability. The ASD sample was limited to fifteen dyads including 14 boys and 1 girl. While this can be attributed to the difference in prevalence of ASD between males and females, a more representative population would have opened up the ability to assess the effects of gender on the results. Furthermore, the sample was skewed towards greater ASD severity, as measured by ADOS Comparison Scores. Within the spectrum of ASD, not all levels of children with ASD were proportionally represented in the sample. All results must be interpreted in the light of a limited sample size. Due to limited time and resources, the TD sample play sessions were not hand-coded for vocalizations. Without the TD dyadic interactions, vocal characteristics between ASD and TD dyads were unable to be compared. These variables would have provided a more complete picture of the sample and would have allowed for more direct comparison between a TD control group and ASD group in terms of specific behavioral measures of communication and their relationship with AF slopes.

While this study was constrained by limitations on time and resources, future studies can expand on the findings of this study. For more detailed analysis on hierarchical temporal clustering, the specific timescales at which AF slopes are diverging can be examined and correlated with vocalization variables. Future directions can explore vocal complexity of the child sample by measuring consonant diversity. This measure would have provided more information about the language level of the ASD sample in addition to an alternate route to determine the relationship between infant language skills and hierarchical temporal clustering. Finally, parent characteristics such as parent education level, SES, and depression, can be run through correlation tests with vocal characteristics and AF slopes. The results would have provided insight into the parent sample, and could have been presented along the existing correlations with the ASD toddler characteristics. Lastly, the limitations of the current study can be addressed to include a larger, more diverse sample, along with an age-matched TD group.

Future studies will be able to provide greater insight beyond the limitations that constrained this current study.

APPENDIX

A. Free Play Session Toys

Toys	Quantity
Puppets	2
Baby Doll	1
Bottles	2
Baby Blanket	1
Knobby Ball	1
Cars	2
Barn	1
Barn Characters	5
Maisy Book	1
Spot Book	1
Curious George Book	1
Little Dog Book	1

B. Vocalization Variable Definitions

Term	Definition
Infant Frequency	Frequency of infant vocalization, i.e. the total number of infant vocalizations
IL Frequency	Frequency of infant linguistic vocalization (IL), i.e. the total number of infant linguistic vocalizations
Parent Frequency	Frequency of parent vocalization, i.e. the total number of parent vocalizations
PT Frequency	Frequency of parent linguistic vocalization, i.e. the total number of parent linguistic vocalizations
Total Parent Infant Frequency	Total Frequency of parent and infant vocalizations, i.e. the total number of infant and parent vocalizations
Total Parent Infant Frequency No SS	Total Frequency of parent and infant vocalizations without simultaneous speech, i.e. the total number of infant and parent vocalizations subtracted by simultaneous speech
Total PT IL Frequency	Total Frequency of parent and infant linguistic vocalizations, i.e. the total number of infant and parent linguistic vocalizations
Total PT IL Frequecy No SS	Total Frequency of parent and infant linguistic vocalizations without simultaneous speech, i.e. the total number of infant and parent linguistic vocalizations subtracted by simultaneous speech
Infant Duration	Average duration of infant vocalization, i.e. the length of the utterance in seconds
IL Duration	Average duration of infant linguistic vocalization, i.e. the length of the utterance in seconds
Parent Duration	Average duration of parent vocalization, i.e. the length of the utterance in seconds
PT Duration	Average duration of parent linguistic vocalization, i.e. the length of the utterance in seconds
Infant Total Duration	Total duration of infant vocalizations, i.e. the sum of all utterance durations
IL Total Duration	Total duration of infant linguistic vocalizations, i.e. the sum of all linguistic utterance durations
Parent Total Duration	Total duration of parent vocalizations, i.e. the sum of all utterance durations
PT Total Duration	Total duration of parent linguistic vocalizations, i.e. the sum of all linguistic utterance durations
Total Parent Infant Duration	Total duration of parent and infant vocalizations, i.e. the sum of all parent and infant vocalization durations
Total Parent Infant Duration No SS	Total duration of parent and infant vocalizations without simultaneous speech, i.e. the sum of all parent and infant vocalization durations subtracted by simultaneous speech duration
Total PT IL Duration	Total duration of parent and infant linguistic vocalizations, i.e. the sum of all parent and infant linguistic vocalization durations
Total PT IL Duration No SS	Total duration of parent and infant linguistic vocalizations without simultaneous speech, i.e. the sum of all parent and infant linguistic vocalization durations subtracted by simultaneous speech duration
Turn	1 Full Turn: speaker 1 vocalization, <3000ms pause, speaker 2 vocalization
Turn Latency	Length of pause between speech with upper boundary of 3000ms
IL to PT Turn	Frequency of turns in this order: infant linguistic vocalization, <3000 ms pause, parent linguistic vocalization
Infant to Parent Turn	Frequency of turns in this order: infant vocalization, <3000 ms pause, parent vocalization
PT to IL Turn	Frequency of turns in this order: parent linguistic vocalization, <3000 ms pause, infant linguistic vocalization
Parent to Infant Turn	Frequency of turns in this order: parent vocalization, <3000 ms pause, infant vocalization
Parent to Parent Turn	Frequency of turns in this order: parent vocalization, <3000 ms pause, parent vocalization
Infant to Infant Turn	Frequency of turns in this order: infant vocalization, <3000 ms pause, infant vocalization
Total Parent Infant Turns	Total turns of interpersonal turns and intrapersonal turns
Total PTIL Turns	Total turns of linguistic interpersonal turns and linguistic intrapersonal turns
VR	Visual Reception age equivalency, Mullen Scales of Early Learning
RL	Receptive language age equivalency, Mullen Scales of Early Learning
EL	Expressive language age equivalency, Mullen Scales of Early Learning
Composite	Language composite of the average of receptive and expressive language
ADOS-2 CS	ADOS Comparison Score, measured on a scale of 1-10 with 10 being the most severe

C. Regression Analysis

A data point has high *leverage* if it has extreme x predictor values, while a point has high *influence* if it has a large impact on the regression analysis (. To determine the effect of outliers, leverage and influence were calculated. A data point with an influence value greater than a Cook's distance of 4/n (n=15) was considered a significant outlier and removed from the dataset.

i. Parent Frequency

Table 7 displays the results from the regression model of parent frequency and AF slope. There was a significant effect of AF slope on parent frequency (*b* = –.52, *t*(13) = -2.19, *p* = .047). This model explained 27% of the variance in frequency of parent vocalizations (R^2 = .27, F(13) = 4.81, p = .047).

Table 7. Regression Model of Parent Frequency With Outlier

Residuals					
	Min	**1Q**	**Median**	**3Q**	**Max**
	-50.305	-18.881	-6.957	24.535	54.729
Coefficients					
	Estimate	**Std. Error**	**t value**	**Pr(>|t|)**	
(Intercept)	301.62	76.09	3.964	0.00162 **	
AF Slope	-221.39	100.95	-2.193	0.04709 *	

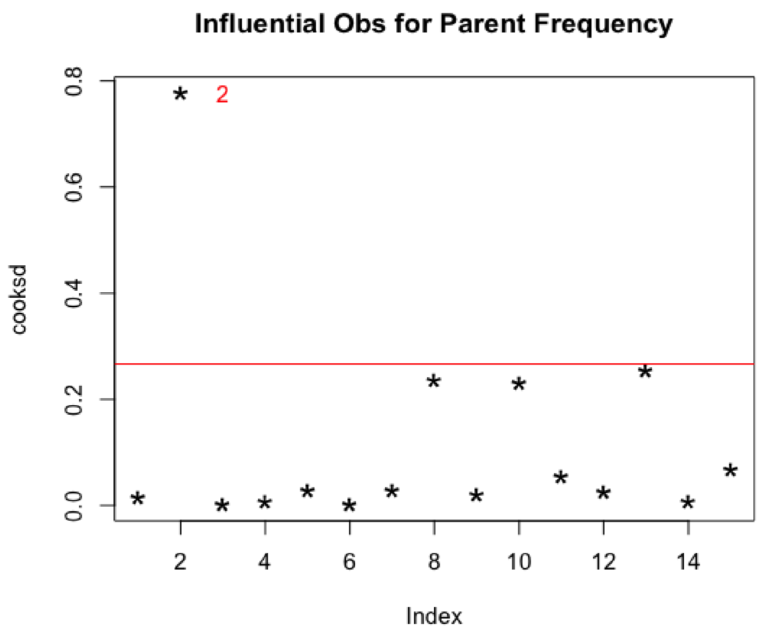
Table 8. Regression Analysis of Parent Frequency Without Outlier

Residuals

	Min	1Q	Median	3Q	Max
	-50.305	-18.881	-6.957	24.535	54.729

Coefficients

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	301.62	76.09	3.964	0.00162 **
AF Slope	-221.39	100.95	-2.193	0.04709 *



ii. Infant-to-Parent Turns

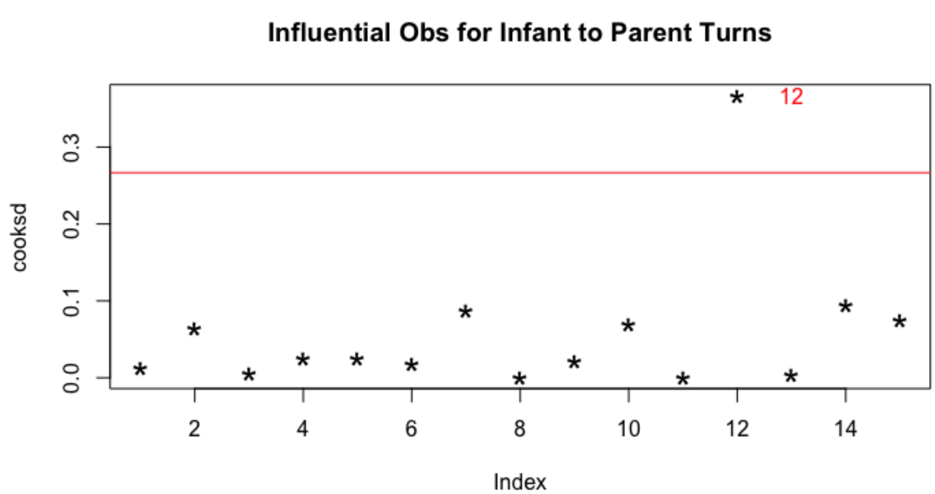
Table 9 and 10 display the results from the regression model. There was a significant effect of AF Slope on infant-to-parent turns (*b* = –.53, *t*(13) = -2.26, *p* = .042). This model explained 28% of the variance in infant-to-parent turns (R^2 = .28, F(13) = 5.11, p = .042). To determine the effect of outliers, leverage and influence were calculated for all data points.

Table 9: Regression Analysis of Infant-to-Parent Turns With Outlier

Residuals					
	Min	**1Q**	**Median**	**3Q**	**Max**
	-22.8289	-6.0273	0.6887	8.1204	17.2197
Coefficients					
	Estimate	**Std. Error**	**t value**	**Pr(>|t|)**	
(Intercept)	87.62	28.38	3.087	0.00866**	
AF Slope	-85.13	37.65	-2.261	0.04154*	

Table 10: Regression Analysis of Infant-to-Parent Turns Without Outlier

Residuals					
	Min	**1Q**	**Median**	**3Q**	**Max**
	-17.714	-7.244	2.471	5.403	15.731
Coefficients					
	Estimate	**Std. Error**	**t value**	**Pr(>|t|)**	
(Intercept)	106.82	25.01	4.272	0.00108 **	
AF Slope	-108.39	32.92	-3.293	0.00643 **	



BIBLIOGRAPHY

Abney, D. H., Paxton, A., Dale, R., & Kello, C. T. (2014). Complexity matching in dyadic conversation. *Journal of Experimental Psychology: General*, *143*(6), 2304–2315. https://doi.org/10.1037/xge0000021

Abney, D. H., Warlaumont, A. S., Oller, D. K., Wallot, S., & Kello, C. T. (2017). Multiple Coordination Patterns in Infant and Adult Vocalizations. *Infancy*, *22*(4), 514–539. https://doi.org/10.1111/infa.12165

Albin, D. D., & Echols, C. H. (1996). Stressed and word-final syllables in infant-directed speech. *Infant Behavior and Development*, *19*(4), 401–418. https://doi.org/10.1016/S0163-6383(96)90002-8

Allen, J. F., Miller, B. W., Ringger, E. K., & Sikorski, T. (1996). *A Robust System for Natural Spoken Dialogue*. http://arxiv.org/abs/cmp-lg/9606023

Anderson, D. K., Lord, C., Risi, S., DiLavore, P. S., Shulman, C., Thurm, A., Welch, K., & Pickles, A. (2007). Patterns of Growth in Verbal Abilities Among Children With Autism Spectrum Disorder. *Journal of Consulting and Clinical Psychology*, *75*(4), 594–604. https://doi.org/10.1037/0022-006X.75.4.594

Autism and Autism Spectrum Disorders. (n.d.). Retrieved April 25, 2021, from https://www.apa.org/topics/autism-spectrum-disorder

Baird Pharr, A. E., Bernstein Ratner, N., & RESCORLA Bryn Mawr College, L. (2000). Syllable structure development of toddlers with expressive specific language impairment. In *Applied Psycholinguistics* (Vol. 21). Cambridge University Press. https://drum.lib.umd.edu/handle/1903/7473

Baranek, G. T. (1999). Autism during infancy: A retrospective video analysis of sensory-motor and social behaviors at 9-12 months of age. *Journal of Autism and Developmental Disorders*, *29*(3), 213–224. https://doi.org/10.1023/A:1023080005650

Boucher, J. (2012). Research review: Structural language in autistic spectrum disorder - Characteristics and causes. In *Journal of Child Psychology and Psychiatry and Allied Disciplines* (Vol. 53, Issue 3, pp. 219–233). J Child Psychol Psychiatry. https://doi.org/10.1111/j.1469-7610.2011.02508.x

Brugman, H., Brugman, H., Russel, A., & Nijmegen, X. (2004). Annotating multi-media / multimodal resources with ELAN. *IN PROCEEDINGS OF LREC*, *46*, 2065--2068. http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.104.1178

Bruner, J. (1983). Play, Thought, and Language. *Peabody Journal of Education*, *60*(3), 60–69. https://doi.org/10.1080/01619568309538407

Chenausky, K., Nelson, C., & Tager-Flusberg, H. (2017). Vocalization rate and consonant production in toddlers at high and low risk for Autism. *Journal of Speech, Language, and Hearing Research*, *60*(4), 865–876. https://doi.org/10.1044/2016_JSLHR-S-15-0400

Cohen, D., Cassel, R. S., Saint-Georges, C., Mahdhaoui, A., Laznik, M.-C., Apicella, F., Muratori, P., Maestro, S., Muratori, F., & Chetouani, M. (2013). Do Parentese Prosody and Fathers' Involvement in Interacting Facilitate Social Interaction in Infants Who Later Develop Autism? *PLoS ONE*, *8*(5), e61402. https://doi.org/10.1371/journal.pone.0061402

Cooper, R. P., & Aslin, R. N. (1990). Preference for Infant-Directed Speech in the First Month after Birth. *Child Development*, *61*(5), 1584. https://doi.org/10.2307/1130766

Delavenne, A., Gratier, M., & Devouche, E. (2013). Expressive timing in infant-directed singing between 3 and 6 months. *Infant Behavior and Development*, *36*(1), 1–13. https://doi.org/10.1016/j.infbeh.2012.10.004

Eilers, R. E., & Oller, D. K. (1994). Infant vocalizations and the early diagnosis of severe hearing impairment. *The Journal of Pediatrics*, *124*(2), 199–203. https://doi.org/10.1016/S0022-3476(94)70303-5

Falk, S., & Kello, C. T. (2017). Hierarchical organization in the temporal structure of infant-direct speech and song. *Cognition*, *163*, 80–86. https://doi.org/10.1016/j.cognition.2017.02.017

Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., Reznick, J. S., & Bates, E. (2007). MacArthur– Bates Communicative Development Inventories: Users guide and technical manual (2nd ed.). Baltimore, MD: Brookes.

Ferguson, C. A. (1964). Baby Talk in Six Languages. *American Anthropologist*, *66*(6_PART2), 103–114. https://doi.org/10.1525/aa.1964.66.suppl_3.02a00060

Fernald, A. (2000). Speech to infants as hyperspeech: Knowledge-driven processes in early word recognition. *Phonetica*, *57*(2–4), 242–254. https://doi.org/10.1159/000028477

Fernald, A., & Mazzie, C. (1991). Prosody and Focus in Speech to Infants and Adults. *Developmental Psychology*, *27*(2), 209–221. https://doi.org/10.1037/0012-1649.27.2.209

Fernald, A., Taeschner, T., Dunn, J., Papousek, M., De Boysson-Bardies, B., & Fukui, I. (1989). A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of Child Language*, *16*(3), 477–501. https://doi.org/10.1017/S0305000900010679

Fisher, C., & Tokura, H. (1996). Acoustic Cues to Grammatical Structure in Infant-Directed Speech: Cross-Linguistic Evidence. *Child Development*, *67*(6), 3192–3218. https://doi.org/10.1111/j.1467-8624.1996.tb01909.x

Franchini, M., Glaser, B., Wood de Wilde, H., Gentaz, E., Eliez, S., & Schaer, M. (2017). Social orienting and joint attention in preschoolers with autism spectrum disorders. *PLOS ONE*, *12*(6), e0178859. https://doi.org/10.1371/journal.pone.0178859

Gamliel, I., Yirmiya, N., Jaffe, D. H., Manor, O., & Sigman, M. (2009). Developmental trajectories in siblings of children with autism: Cognition and language from 4 months to 7 years. *Journal of Autism and Developmental Disorders*, *39*(8), 1131–1144. https://doi.org/10.1007/s10803-009-0727-2

Gernsbacher, M. A., Geye, H. M., & Weismer, S. E. (2005). *The role of language and communication impairments within autism*. https://doi.org/10.1075/tilar.4.06ger

Girolametto, L. E. (1988). Improving the social-conversational skulls of developmentally delayed children: An intervention study. *Journal of Speech and Hearing Disorders*, *53*(2), 156–167. https://doi.org/10.1044/jshd.5302.156

Goldstein, M. H., & Schwade, J. A. (2008). Social feedback to infants' babbling facilitates rapid phonological learning. *Psychological Science*, *19*(5), 515–523. https://doi.org/10.1111/j.1467-9280.2008.02117.x

Golinkoff, R. M., Can, D. D., Soderstrom, M., & Hirsh-Pasek, K. (2015). (Baby)Talk to Me. *Current Directions in Psychological Science*, *24*(5), 339–344. https://doi.org/10.1177/0963721415595345

Greenwood, C. R., Walker, D., & Buzhardt, J. (2010). The Early Communication Indicator for Infants and Toddlers. *Journal of Early Intervention*, *32*(5), 310–334. https://doi.org/10.1177/1053815110392335

Hilbrink, E. E., Gattis, M., & Levinson, S. C. (2015). Early developmental changes in the timing of turn-taking: a longitudinal study of mother–infant interaction. *Frontiers in Psychology*, *6*, 1492. https://doi.org/10.3389/fpsyg.2015.01492

Jaffe, J., Beebe, B., Feldstein, S., Crown, C. L., & Jasnow, M. D. (2001). Rhythms of dialogue in infancy: coordinated timing in development. *Monographs of the Society for Research in Child Development*, *66*(2), i–viii, 1–132.

Kaplan, P. S., Goldstein, M. H., Huckeby, E. R., Owren, M. J., & Cooper, R. P. (1995). Dishabituation of visual attention by infant- versus adult-directed speech: Effects of frequency modulation and spectral composition. *Infant Behavior and Development*, *18*(2), 209–223. https://doi.org/10.1016/0163-6383(95)90050-0

Katz, G. S., Cohn, J. F., & Moore, C. A. (1996). A Combination of Vocal f0 Dynamic and Summary Features Discriminates between Three Pragmatic Categories of Infant-Directed Speech. *Child Development*, *67*(1), 205–217. https://doi.org/10.1111/j.1467-8624.1996.tb01729.x

Kemler Nelson, D. G., Hirsh-Pasek, K., Jusczyk, P. W., & Cassidy, K. W. (1989). How the prosodic cues in motherese might assist language learning. *Journal of Child Language, 16*(1), 55–68. https://doi.org/10.1017/S030500090001343X

Kuhl, P. K., Andruski, J. E., Chistovich, I. A., Chistovich, L. A., Kozhevnikova, E. V., Ryskina, V. L., Stolyarova, E. I., Sundberg, U., & Lacerda, F. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science, 277*(5326), 684–686. https://doi.org/10.1126/science.277.5326.684

Lahey, M., & Bloom, L. (1988). *Language Disorders and Language Development* (1st ed.). Macmillan.

Lohmander, A., Holm, K., Eriksson, S., & Lieberman, M. (2017). Observation method identifies that a lack of canonical babbling can indicate future speech and language problems. *Acta Paediatrica, International Journal of Paediatrics*, *106*(6), 935–943. https://doi.org/10.1111/apa.13816

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

Luque, J., Luque, B., & Lacasa, L. (2015). Scaling and universality in the human voice. *Journal of the Royal Society Interface*, *12*(105). https://doi.org/10.1098/rsif.2014.1344

McDaniel, J., Yoder, P., Estes, A., & Rogers, S. J. (2020). Predicting expressive language from early vocalizations in young children with autism spectrum disorder: Which vocal measure is best? *Journal of Speech, Language, and Hearing Research*, *63*(5), 1509–1520. https://doi.org/10.1044/2020_JSLHR-19-00281

McDaniel, J., Yoder, P., Estes, A., & Rogers, S. J. (2020). Validity of Vocal Communication and Vocal Complexity in Young Children with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, *50*(1), 224–237. https://doi.org/10.1007/s10803-019-04248-x

Mullen, E. M. (1995). *Mullen Scales of Early Learning* (AGS ed.). Circle Pines, MN: American Guidance Service Inc.

Oller, D. K., Niyogi, P., Gray, S., Richards, J. A., Gilkerson, J., Xu, D., Yapanel, U., & Warren, S. F. (2010). Automated vocal analysis of naturalistic recordings from children with autism, language delay, and typical development. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(30), 13354–13359. https://doi.org/10.1073/pnas.1003882107

Oller, D. K. (2001). The Emergence of Speech Capacity. *The Journal of the Acoustical Society of America*, *110*(3), 1237–1238. https://doi.org/10.1121/1.1388001

Oller, D. K. (1973). The effect of position in utterance on speech segment duration in English. *The Journal of the Acoustical Society of America*, *54*(5), 1235–1247. https://doi.org/10.1121/1.1914393

Oller, D. K., Eilers, R. E., Neal, A. R., & Cobo-Lewis, A. B. (1998). Late onset canonical babbling: A possible early marker of abnormal development. *American Journal on Mental Retardation*, *103*(3), 249–263. https://doi.org/10.1352/0895-8017(1998)103<0249:LOCBAP>2.0.CO;2

Oller, D. K., Eilers, R. E., Neal, A. R., & Schwartz, H. K. (1999). Precursors to speech in infancy: The prediction of speech and language disorders. *Journal of Communication Disorders*, *32*(4), 223–245. https://doi.org/10.1016/S0021-9924(99)00013-1

Osterling, J. A., Dawson, G., & Munson, J. A. (2002). Early recognition of 1-year-old infants with autism spectrum disorder versus mental retardation. *Development and Psychopathology*, *14*(2), 239–251. https://doi.org/10.1017/s0954579402002031

Osterling, J., & Dawson, G. (1994). Early recognition of children with autism: A study of first birthday home videotapes. *Journal of Autism and Developmental Disorders*, *24*(3), 247–257. https://doi.org/10.1007/BF02172225

Ozonoff, S., Iosif, A. M., Baguio, F., Cook, I. C., Hill, M. M., Hutman, T., Rogers, S. J., Rozga, A., Sangha, S., Sigman, M., Steinfeld, M. B., & Young, G. S. (2010). A Prospective Study of the Emergence of Early Behavioral Signs of Autism. *Journal of the American Academy of Child and Adolescent Psychiatry*, *49*(3), 256–266. https://doi.org/10.1097/00004583-201003000-00009

Patten, E., Belardi, K., Baranek, G. T., Watson, L. R., Labban, J. D., & Oller, D. K. (2014). Vocal patterns in infants with autism spectrum disorder: Canonical babbling status and vocalization frequency. *Journal of Autism and Developmental Disorders*, *44*(10), 2413–2428. https://doi.org/10.1007/s10803-014-2047-4

Paul, R., & Jennings, P. (1992). Phonological behavior in toddlers with slow expressive language development. *Journal of Speech and Hearing Research*, *35*(1), 99–107. https://doi.org/10.1044/jshr.3501.99

Paul, R., Fuerst, Y., Ramsay, G., Chawarska, K., & Klin, A. (2011). Out of the mouths of babes: Vocal production in infant siblings of children with ASD. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *52*(5), 588–598. https://doi.org/10.1111/j.1469-7610.2010.02332.x

Pegg, J. E., Werker, J. F., & McLeod, P. J. (1992). Preference for infant-directed over adult-directed speech: Evidence from 7-week-old infants. *Infant Behavior and Development*, *15*(3), 325–345. https://doi.org/10.1016/0163-6383(92)80003-D

Peppé, S., McCann, J., Gibbon, F., O'Hare, A., & Rutherford, M. (2007). Receptive and expressive prosodic ability in children with high-functioning autism. *Journal of Speech, Language, and Hearing Research*, *50*(4), 1015–1028. https://doi.org/10.1044/1092-4388(2007/071)

Pharr, A. B., Ratner, N. B., & Rescorla, L. (2000). Syllable structure development of toddlers with expressive specific language impairment. *Applied Psycholinguistics*, *21*(4), 429–449. https://doi.org/10.1017/S014271640000401X

Plumb, A. M., & Wetherby, A. M. (2013). Vocalization development in toddlers with autism spectrum disorder. *Journal of Speech, Language, and Hearing Research*, *56*(2), 721–734. https://doi.org/10.1044/1092-4388(2012/11-0104)

Räsänen, O., Kakouros, S., & Soderstrom, M. (2018). Is infant-directed speech interesting because it is surprising? – Linking properties of IDS to statistical learning and attention at the prosodic level. *Cognition, 178*, 193–206. https://doi.org/10.1016/j.cognition.2018.05.015

Ratner, N. B. (1986). Durational cues which mark clause boundaries in mother–child speech. *Journal of Phonetics*, *14*(2), 303–309. https://doi.org/10.1016/s0095-4470(19)30670-9

Rescoria, L., & Ratner, N. B. (1996). Phonetic profiles of toddlers with specific expressive language impairment (SLI-E). *Journal of Speech, Language, and Hearing Research*, *39*(1), 153–165. https://doi.org/10.1044/jshr.3901.153

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926–1928. https://doi.org/10.1126/science.274.5294.1926

Santarcangelo, S., & Dyer, K. (1988). Prosodic aspects of motherese: Effects on gaze and responsiveness in developmentally disabled children. *Journal of Experimental Child Psychology*, *46*(3), 406–418. https://doi.org/10.1016/0022-0965(88)90069-0

Schoen, E., Paul, R., & Chawarska, K. (2011). Phonology and vocal behavior in toddlers with autism spectrum disorders. *Autism Research*, *4*(3), 177–188. https://doi.org/10.1002/aur.183

Seidl, A., & Johnson, E. K. (2006). Infant word segmentation revisited: Edge alignment facilitates target extraction. *Developmental Science*, *9*(6), 565–573. https://doi.org/10.1111/j.1467-7687.2006.00534.x

Sheinkopf, S. J., Iverson, J. M., Rinaldi, M. L., & Lester, B. M. (2012). Atypical Cry Acoustics in 6-Month-Old Infants at Risk for Autism Spectrum Disorder. *Autism Research*, *5*(5), 331–339. https://doi.org/10.1002/aur.1244

Sheinkopf, S. J., Mundy, P., Oller, D. K., & Steffens, M. (2000). Vocal atypicalities of preverbal autistic children. *Journal of Autism and Developmental Disorders*, *30*(4), 345–354. https://doi.org/10.1023/A:1005531501155

Shukla, M., White, K. S., & Aslin, R. N. (2011). Prosody guides the rapid mapping of auditory word forms onto visual objects in 6-mo-old infants. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(15), 6038–6043. https://doi.org/10.1073/pnas.1017617108

Snow, C. E. (2012). Beginning from Baby Talk: twenty years of research on input in interaction. In *Input and Interaction in Language Acquisition* (pp. 3–12). Cambridge University Press. https://doi.org/10.1017/cbo9780511620690.002

Spinelli, M., Fasolo, M., & Mesman, J. (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

Spotlight On: Delay Between First Concern to Accessing Services | Autism | NCBDDD | CDC. (n.d.). Retrieved April 25, 2021, from https://www.cdc.gov/ncbddd/autism/addm-community-report/delay-to-accessing-services.html

Stern, D. N., Spieker, S., Barnett, R., & Mackain, K. (1983). The Prosody of Maternal Speech: Infant Age and Context Related Changes. *Journal of Child Language*, *10*(1), 1–15. https://doi.org/10.1017/S0305000900005092

Stern, D. N. (1985). *The Interpersonal World of the Infant: A View from Psychoanalysis and Developmental Psychology*. Basic Books. https://www.pep-web.org/document.php?id=zbk.016.0001a

Talbott, M. R., Nelson, C. A., & Tager-Flusberg, H. (2016). Maternal Vocal Feedback to 9-Month-Old Infant Siblings of Children with ASD. *Autism Research*, *9*(4), 460–470. https://doi.org/10.1002/aur.1521

Tamis-LeMonda, C. S., Bornstein, M. H., & Baumwell, L. (2001). Maternal responsiveness and children's achievement of language milestones. *Child Development*, *72*(3), 748–767. https://doi.org/10.1111/1467-8624.00313

Thiessen, E. D., Hill, E. A., & Saffran, J. R. (2005). Infant-Directed Speech Facilitates Word Segmentation. *Infancy*, *7*(1), 53–71. https://doi.org/10.1207/s15327078in0701_5

Trainor, L. J., & Adams, B. (2000). Infants' and adults' use of duration and intensity cues in the segmentation of tone patterns. *Perception and Psychophysics*, *62*(2), 333–340. https://doi.org/10.3758/BF03205553

Trainor, L. J., & Desjardins, R. N. (2002). Pitch characteristics of infant-directed speech affect infants' ability to discriminate vowels. *Psychonomic Bulletin and Review*, *9*(2), 335–340. https://doi.org/10.3758/BF03196290

Trehub, S. E., Unyk, A. M., & Trainor, L. J. (1993). Maternal singing in cross-cultural perspective. *Infant Behavior and Development*, *16*(3), 285–295. https://doi.org/10.1016/0163-6383(93)80036-8

Vouloumanos, A., & Curtin, S. (2014). Foundational Tuning: How Infants' Attention to Speech Predicts Language Development. *Cognitive Science*, *38*(8), 1675–1686. https://doi.org/10.1111/cogs.12128

Wang, Y., Lee, C. S., & Houston, D. M. (2016). Infant-directed speech reduces English-learning infants' preference for trochaic words. *The Journal of the Acoustical Society of America*, *140*(6), 4101–4110. https://doi.org/10.1121/1.4968793

Warlaumont, A. S., Richards, J. A., Gilkerson, J., & Oller, D. K. (2014). A Social Feedback Loop for Speech Development and Its Reduction in Autism. *Psychological Science*, *25*(7), 1314–1324. https://doi.org/10.1177/0956797614531023

Warren, S. F., Brady, N., Sterling, A., Fleming, K., & Marquis, J. (2010). Maternal responsivity predicts language development in young children with fragile X syndrome. *American Journal on Intellectual and Developmental Disabilities*, *115*(1), 54–75. https://doi.org/10.1352/1944-7558-115.1.54

Weisleder, A., & Fernald, A. (2013). Talking to Children Matters: Early Language Experience Strengthens Processing and Builds Vocabulary. *Psychological Science*, *24*(11), 2143–2152. https://doi.org/10.1177/0956797613488145

Werker, J. F., Pegg, J. E., & McLeod, P. J. (1994). A cross-language investigation of infant preference for infant-directed communication. *Infant Behavior and Development*, *17*(3), 323–333. https://doi.org/10.1016/0163-6383(94)90012-4

Yoder, P. J., & Warren, S. F. (1999). Maternal Responsivity Mediates the Relationship Between Prelinguistic Intentional Communication and Later Language. *Journal of Early Intervention*, *22*(2), 126–136. https://doi.org/10.1177/105381519902200205

Zwaigenbaum, L., Bryson, S., Rogers, T., Roberts, W., Brian, J., & Szatmari, P. (2005). Behavioral manifestations of autism in the first year of life. *International Journal of Developmental Neuroscience*, *23*(2-3 SPEC. ISS.), 143–152. https://doi.org/10.1016/j.ijdevneu.2004.05.001