Dyadic Coordination of Vocalizations and Pauses in Autistic Children and their Caregivers

Eun Ae Choi

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2023

Reading Committee:

Sara T. Kover, Chair

Amy E. Pace

Wendy L. Stone

Program Authorized to Offer Degree:

Speech and Hearing Sciences

University of Washington

Abstract

Dyadic Coordination of Vocalizations and Pauses in Autistic Children and their Caregivers

Eun Ae Choi

Chair of the Supervisory Committee:

Sara T. Kover

Department of Speech and Hearing Sciences

There is emerging evidence that infants with autism spectrum disorder (ASD) demonstrate differences in dyadic coordination of vocalization and pauses with their caregivers throughout infancy and childhood compared to neurotypical children (Northrup et al, 2015; Seidl et al., 2018; Warlaumont et al., 2014). In fact, reduced dyadic coordination of vocalization characterized as larger differences in latency to respond to one another (i.e., latency difference) and overlapping vocalizations were observed to be associated with later receptive and expressive language delays in neurotypical and neurodiverse infant-caregiver dyads (Northrup & Iverson, 2015). However, dyadic coordination and its relationship to language in older, school-age autistic children are poorly understood. The current study investigated dyadic coordination variables (child and caregiver's percent overlapping speech, latency difference) and conversational state variables (frequency and mean duration of vocalizations and pauses) in dyads of autistic or typically developing school-age children with their caregivers engaged in parent-child-play sessions in the home environment. Results showed that autistic children's vocalization and within-speaker pauses (i.e., pauses between vocalizations made by the same speaker) were significantly less compared to typically developing children. No group differences were observed in caregivers' percent overlapping speech or conversational state variables. One dyadic coordination variable, latency difference, significantly and uniquely predicted concurrent receptive vocabulary size, but was not a reliable predictor of composite language scores. Contrary to previous findings, latency difference was positively associated with concurrent language abilities in school-age children. These findings suggest that dyadic coordination variables present differently in older school-age children engaged in parent-child play sessions compared to infant-caregiver dyads or dyads engaged in structured tasks.

Dedication

I dedicate this to my dad, Jin Ho Choi.

For your lessons on perseverance and tenacity on the ski slopes.

Acknowledgements

I would like to express my deepest gratitude to my advisor, Dr. Sara Kover. Her constant and unwavering support kept me rooted throughout the entire doctoral program. I am forever thankful for the hours of meetings, document revisions, and letters of supports that she generously gave to ensure my success as a doctoral student and as a budding researcher. Her passion in research, service to the department, and dedication to each of her students are truly inspirational.

I am also very grateful for my doctoral committee members, Dr. Amy Pace, Dr. Wendy Stone, and Dr. Ilene Schwartz. Their guidance and support deepened my understanding in the field in ways I could not have imagined. Thank you for sharing your wisdom and widening my perspective. A special thanks to Dr. Andrew Brown for his service as the doctoral student advisor. His commitment in connecting with and advocating for each doctoral student has left a mark in many of us, and I hope to model after him in my future advisory roles.

I am immensely thankful for all the NeudLL lab members and the families who gave their time to participate in our studies. To Emily Fowler, thank you for being my lab mate, colleague, and friend. Your warm presence, kindness and delicious snacks brought all of us together. Thank you for creating that safe space and community in the lab.

I am thankful for my husband and daughter for their patience and kindness through every step of my study. Thank you, Wilson and Hailey, for celebrating all the small successes and milestones along the way, giving me strength to press on and reach my goals. I would like to also thank my mom, dad, sister, and my mother-in-law for taking care of us. This dissertation would not have been possible without all your love, meals, prayers, and support.

Table of Contents

DYADIC PROCESSES

From a very young age, children and their caregivers engage in coordinated "dance-like" interactions with each other across various modalities. Typically developing 1-year-olds with their mothers engage in locomotor synchrony, matching the spatial and temporal aspects of their movements (Hoch, Ossmy, Cole, Hasan, & Adolph, 2021). Typically developing 6-month-olds demonstrate mutual and coordinated eye-gaze with their caregivers (Northrup & Iverson, 2020; Northrup & Iverson, 2020). Most remarkably, infants as young as 3 to 4 months of age engage in a rhythmic "dialogue" with their caregivers, mimicking the coordinated turn-taking structures of adult conversations (Gratier et al., 2015; Jaffe, Beebe, Feldstein, Crown, & Jasnow, 2001;Northrup & Iverson, 2020).

From an anthropological perspective, these "dance-like" interactions between dyads are theorized to play a critical role in the infants' well-being and long-term outcomes. For instance, synchronous interactions within the dyads have been observed to improve infants' homeostatic regulation (e.g., feeding, sleeping, arousal), as well as facilitate secure attachments between the child with his caregiver, thus, potentially securing the chances of the infant's survival (for review; Harrist & Waugh, 2002).

From a developmental perspective, it is hypothesized that "… turn-taking, a key characteristic of dyadic synchrony [coordination], not only gives children practice in the give-and-take of social conversations…, but that 'two person exchange rituals' – repeated, structured, reciprocal activities—lead to the development of shared meaning… and therefore, language learning" (Harrist & Waugh, 2002, p. 572). In other words, the back-and-forth coordination of vocalization creates an organized structure in an interaction, which in turn allows for the developing infant to understand and practice shared meanings, such as language. As such, dyadic processes between infants and their caregivers have been studied across disciplines (e.g., developmental psychology, psycholinguistics, communication sciences and disorders) at an increasing rate throughout the past four decades (Provenzi, di Minico, Giusti, Guida, & Müller, 2018).

The study of dyadic processes between children and their caregivers has theoretical and clinical implications in the development of cognitive, social, and linguistic domains. For instance, certain dyadic processes observed during the early stages of development are associated with the child's later outcomes across biopsychosocial domains, including language (Feldman, 2010; Harrist & Waugh, 2002; Hove & Risen, 2009; Jaffe et al., 2001; Leclère et al., 2014; Provenzi et al., 2018). The growing evidence on the role of early dyadic processes serves as a critical foundation to theories on how children acquire and master complex abilities. Moreover, the differences in quality or quantity of dyadic processes may potentially serve as early markers of neurodevelopmental differences in certain populations and contribute to the process of diagnosis and treatment support for the families (Feldman, 2007; Feldman, Greenbaum, & Yirmiya, 1999; Kellerman et al., 2020; Leclère et al., 2014; Nomikou, Koke, & Rohlfing, 2017; Rauchbauer & Grosbras, 2020; Skuban, Shaw, Gardner, Supplee, & Nichols, 2006; Trevarthen & Daniel, 2005; Wan et al., 2013).

Given the steady growth in the study of dyadic processes in children and their caregivers over the past years, researchers have sought to investigate recurring concepts and themes in the field. Provenzi, di Minico, Giusti, Guida, and Müller (2018) constructed a useful framework to capture themes in this area of research based on their systematic review of 82 studies across disciplines, examining the dyadic processes between children and their caregivers. They presented a framework that serves as one way to organize various types of dyadic processes and identify the operationalization of those concepts across studies. Among the many concepts highlighted by Provenzi et al. (2018), there are two theoretical concepts (i.e., reciprocity and mutuality) that are better defined as "global meta-theoretical concepts" and appeared to drive the broader direction of the studies in the systematic review. The meta-theoretical concept "reciprocity" assumes that each member of the dyad contributes *equally* in the relevance, intensity, and frequency of their interactions. For instance, in a study that examined the effects of infant irritability on caregiver–infant interactions 3 months later, reciprocity was defined as both the caregiver and the infant engaging in the same behavioral mode (e.g., touch, eye-gaze, vocalization) at a specific time interval (Lowinger, 1999). Newborn infants who demonstrated higher level of irritability at 3 days, were observed to engage in fewer touch reciprocal behaviors (e.g., holding/being held) but more social reciprocal behaviors (e.g., smiling, vocalization, eye-gaze) during the 10-week observation. The underlying assumption in the study was that reciprocity meant equal contribution in timing and mode from the infant and the caregiver.

On the other hand, the meta-theoretical concept of "mutuality" assumes that while both the child and the caregiver contribute to the dyadic system, their contributions may differ in quality or quantity. For instance, in a study that examined the relationship between caregivers' and infants' vocalization frequency, caregivers by far, produced higher frequency of vocalizations than infants (Roe & Drivas, 1997). However, the authors observed that infants in one group (i.e., moderate mother's vocalization) engaged in the highest number of vocalization compared to the infants in the comparison groups (i.e., high/low mother's vocalization). Hence, the authors concluded that a moderate frequency of mother's vocalization appeared to be the most balanced interaction, optimal for encouraging spontaneous and independent infant vocalizations. In sum, the two meta-theoretical concepts play an important role in determining the overall nature of studies on dyadic processes (e.g., interpretation of results, operationalization of behavioral data) and introduce the idea that fluent or optimal dyadic coordination can be achieved through both reciprocal and mutual interactions.

Under the umbrella of two global meta-theoretical concepts (reciprocity and mutuality), seven dyadic processes are interconnected in the dynamic framework (See Figure 1). The seven specific constructs (i.e., contingency, coordination, matching, reparation, attunement, mirroring, and synchrony) are dyadic *processes* that have both theoretical and methodological definitions (e.g., proportion, frequency, or duration of certain behaviors or correlation indices). Provenzi et al. (2018) emphasizes that some processes are considered low-level, meaning it usually occurs earlier in development and involves minimal understanding of intentions or the psychological state of the communication partner. The high-level processes are built on the low-level processes and require more advanced skills in recognizing inner states of the communication partners and also the ability to monitor and change one's behavior across time.

Contingency, as defined by Provenzi et al. (2018), is a low-level dyadic process that emerges very early in development (Murray et al., 2016); it is the reciprocal adjustment of behaviors, such as eye-gaze or vocalizations, between dyads that occur within a short time frame. For instance, the caregiver vocalizing within 2 to 3 seconds after the child's vocalization is one example of a contingent process. Repeated prolonged states of contingent moments between the dyads lead to the dyadic process called *coordination*; a timed, rhythmic, and bidirectional interaction that includes initiation and response of behaviors from both individuals of a dyad. This ability to coordinate behaviors is a critical foundation to understanding intentions, anticipation of affects (attunement and mirroring), and finally the ability to engage in a smooth synchronous interaction. Dyadic synchrony encompasses the ability to coordinate not just behaviors, but also psychological and inner-state forms through reiterations of matching and unmatching of behaviors in a smooth manner (Leclère et al., 2014; Provenzi et al., 2018). The dynamic interconnected nature of the dyadic processes portrayed in the framework proposed by Provenzi et al. (2018) displays the complexity of human interaction and the compounding effects of skills across the child's development.

Nevertheless, there are gaps in the literature that examine dyadic processes especially in the natural environment and among populations that may benefit from deeper understanding of the effects of dyadic processes on later developmental outcomes, such as language abilities. In the systematic review by Provenzi et al. (2018), roughly 70% of the articles examined dyadic processes in the lab environment, whereas only 30% of the studies were conducted in the natural home environment. Moreover, despite the potential clinical implications of dyadic processes in neurodiverse individuals (Wan et al., 2013), most of the studies included in Provenzi et al (2018) were on typically developing caregiver-infant dyads. In another systematic review on synchrony in mother-child interaction, only approximately 20% of their selected studies investigated infants with developmental impairments or psychopathologies (Leclère et al., 2014). As such, dyadic processes in the context of natural home environments in populations with different or delayed developmental trajectories warrants further investigation.

THE DYADIC COORDINATION OF VOCALIZATIONS AND PAUSES

There is emerging evidence that a specific type of coordination may be linked to later developmental outcomes in children with typical development or neurodiverse individuals. The dyadic coordination of vocalization and pauses (i.e., the gaps between vocalizations) is a phenomenon observed beginning in infancy and is hypothesized to be one of the foundational building blocks that facilitates development across the cognitive and social domains (Feldman, 2007; Harrist & Waugh, 2002). The coordination of vocalization and pauses is unique because it exists before the development of spoken words, but resembles the structure of a mature adult-conversation, hence, sometimes called "proto-conversations" (Donnelly & Kidd, 2021; Gratier et al., 2015). By coding pauses and vocalizations that occur between child-caregiver dyads, researchers have revealed that in general, the back-and-forth coordination of vocalization remains fairly consistent in terms of timing and duration throughout the early stages of typical development (Beebe, Alson, Jaffe, Feldstein, & Crown, 1988; Gratier et al., 2015; Northrup & Iverson, 2020).

From 3 months to 12 months of age, dyads consistently demonstrate the least likelihood to vocalize over the other's vocalization (i.e., to produce simultaneous speech), whereas the maximum likelihood of vocalization peaks for both infant and caregiver at approximately 2.5 seconds after the vocalization of their communication partners (Northrup & Iverson, 2020). This observation is a stark comparison to the coordination of mutual eye-gaze: unlike the consistent patterns observed in the coordination of vocalization throughout the first year of development, typically developing infants at 3 months of age steadily progress through irregular patterns and coordination of eye-gaze to increased coordination and organization across the first year of development. This comparison between coordination of vocalization versus the coordination of eye-gazes accentuates the constant and persistent nature of the coordination of vocalization across the first year of typical development.

Similar global patterns in the coordination of vocalization have also been observed in dyads across different cultures (Farran, Yoo, Lee, Bowman, & Oller, 2019; Gratier, 2003). Farran et al. (2019) observed that Lebanese Arabic-speaking mother-infant dyads demonstrated similar global patterns of vocalization (e.g., back-and-forth, rhythmic exchanges) with subtle differences in the pattern of infant vocalization compared to American English-speaking dyads. The authors hypothesized that the Lebanese infants clustering most of their vocalizations to the first 0.5 seconds after their mother's vocalizations may be due to the phonotactic structures of the languages and potentially due to different child-rearing practices across cultures. Global similarities with subtle differences in timing of vocalization were also observed across dyads from India, France, and the United States (Gratier, 2003). Most notably, coordination of vocalization in immigrant Indian dyads living in the United States were observed to assimilate to American dyads (Gratier, 2003). As such, evidence suggests that the coordination of vocalization between infant-caregiver dyads is a universal phenomenon with subtle differences in timing across dyads from different cultures and languages.

A closer examination of the coordination of vocalization and pauses reveal that different patterns in dyadic vocalization may be uniquely associated with various social and cognitive outcomes in typically developing infants. In a hallmark study, Jaffe, Beebe, Feldstein, Crown, and Jasnow (2001) observed patterns of vocalization and pauses between 4-month-old typically developing infant-adult dyads. The authors varied the experiment conditions by infants' familiarity of the communication partners (e.g., caregiver vs. stranger) and the environment (e.g., home vs. laboratory). Further, they analyzed the predictability of the infants' attachment level and cognitive skills at 12 months of age using multiple variables that examined the bidirectional nature of vocalizations and pauses between dyads. General cognition was measured using the

Bayley Scales of Infant Development (Bayley, 1993), which is correlated with later language and reading abilities at ages 6 and 8 years (Siegel, 1989). The dyads in the stranger-laboratory condition displayed an increase in pauses and decrease in duration of vocalization turns compared to the other conditions (e.g., infant-mother-laboratory, infant-stranger-home). In addition, infants paired with strangers in the laboratory environment demonstrated the highest degree of bidirectional coordination of vocalization behaviors (i.e., vocalization behavior of one speaker is highly correlated and predictive of the other speaker's vocalization behavior) compared to dyads in the other conditions. Jaffe et al. (2001) observed that midrange degree of coordination between dyads when the infants were 4 months old predicted secure attachments at 12 months, whereas low and high degree of coordination predicted insecure attachment. In contrast, high degree of coordination, observed in the infant-stranger-laboratory condition, predicted higher scores on the test for cognition at 12 months. The authors speculated that the dyads were engaging in a high degree of coordination to increase their sense of control and predictability in conditions that lacked familiarity and certainty. The timing and duration of vocalizations and pauses between dyads studied in Jaffe et al. (2001) revealed that coordination of vocalizations is a dynamic variable that depends on contextual factors such as the communication partner and the environment. Moreover, dyadic coordination variables observed in different contexts appear to uniquely predict various outcomes at later stages of typical development.

Similar observations were made in a study that examined the degree of coordination between older children with speech differences and their caregivers (Savelkoul, Zebrowski, Feldstein, & Cole-Harding, 2007). Vocalization coordination between 4- to 9-year-old children who stutter and their caregivers in their home environment reflected more mutual influence or accommodation compared to the interaction between 8- to 9-year-old typically developing child-caregiver dyads, (Savelkoul et al., 2007). More specifically, when caregivers of children who stutter interrupted their child (i.e., interruptive simultaneous speech), the child was more likely to interrupt in the following turn for a similar amount of time, and vice versa. The authors speculated that the dyads of children who stutter and their caregivers may demonstrate heightened sensitivity to the temporal characteristics of their conversation partners, potentially due to unpredictability caused by stuttering behaviors. This evidence suggests that the different styles and patterns of coordination of vocalization may not necessarily be an indicator of deficit, but rather an indicator of adaptation and responsiveness within the dyads. The authors urged future researchers to investigate other potential factors, such as the child's social skills, to better understand the internal and external factors that influence the variabilities in the coordination of vocalization and pauses between dyads.

The coordination of vocalization is also present in populations with differences in cognitive abilities. Jasnow et al. (1988) observed that the degree of coordination in vocalization behaviors between infants with Down syndrome, a chromosomal disorder, and their caregivers demonstrated a lower level of coordination (i.e., vocalization behaviors of caregivers were less correlated with vocalization behaviors of infants and vice versa) compared to typically developing infant dyads at 4 months. However, the difference in level of coordination was no longer significant at 9 months of age. The authors suggested that dyadic coordination of vocalization may be an independent factor from cognitive abilities, one of the core deficits of individuals with Down syndrome. The relationship between nonverbal cognition and coordination of vocalization across development or in neurodiverse populations remains unknown; however, it is apparent that variability in coordination of vocalization is worthwhile to explore in its own right and does not appear to be reliant on advanced cognitive abilities.

Thus far, the coordination of vocalization and pauses appears to be a universal and global phenomenon and certain characteristics of coordination, such as the bidirectional or predictive nature of the vocalization and pauses between dyads, are associated with later skills in typical development (Farran et al., 2019; Jaffe et al., 2001). There is also evidence that the duration of pauses within (intrapersonal) and between speakers may be unique aspects of coordination worth investigating. Studies have found congruence and matching in the duration of between speaker pauses in the course of vocal interaction between infants and their mothers (Beebe et al., 1988), as well as in young school-aged children interacting with their peers (Welkowitz, Cariffe, & Feldstein, 1976). The matching of between speaker pauses appear to be strongly correlated with the perception of warmth of their communication partners in typically developing adult dyads (Welkowitz & Kuc, 1973). Some studies propose that lack of congruency in the duration of between speaker pauses may be indicative of reluctance to participate or difficulties in social skills (Feldstein, Konstantareas, Oxman, & Webster, 1982) and potentially predictive of reduced later language abilities (Northrup & Iverson, 2015).

In summary, micro-coded analysis of vocalization and pauses (i.e., precise characterization of frequency and duration) between typically developing infants and their caregivers have revealed that back-and-forth “proto-conversations” are consistent across the first year of typical development (Gratier et al., 2015; Jaffe et al., 2001; Northrup & Iverson, 2020) and cross-culturally (Farran et al., 2019; Gratier, 2003). The degree of coordination between dyads changes according to the familiarity of the infant’s communication partner and environment and is predictive of later social and cognitive skills (Jaffe et al., 2001). In fact, higher degree of coordination between dyads may be reflective of their attempt to bring a sense of control and predictability in the presence of unfamiliarity, as found in adults who interact with children who stutter (Savelkoul et al., 2007). Also, the duration of between-speaker pauses appears to be a unique aspect of dyadic coordination associated with social factors in mother-infant and adult peer dyads (Beebe et al., 1988; Welkowitz & Kuc, 1973), and potentially a marker or index for social skills (Feldstein et al., 1982; Welkowitz et al., 1976) or later language abilities (Northrup & Iverson, 2015) in typically developing or neurodiverse populations. Lastly, in some populations of children with developmental differences, such as children with Down syndrome, dyadic coordination of vocalization and pauses may appear reduced during earlier stages of development, but may be observed to be comparable to typically developing children later in development (Jasnow et al., 1988).

Dyadic coordination of vocalization in autistic children

A neurodiverse individual diagnosed with autism spectrum disorder (ASD) may display a combination or wide spectrum of abilities in social skills, verbal, or nonverbal communication, and display various restrictive repetitive behaviors (American Psychiatric Association [APA], 2013). Micro-coding the coordination of vocalization and pauses between autistic individuals and their conversation partners may shed light on how variability in dyadic coordination relates to concurrent abilities or later outcomes. Investigating factors that contribute to positive outcomes may be useful in developing treatment targets and support systems for autistic individuals. For instance, language abilities have been strongly associated with health-related quality of life and behaviors of concern (e.g., aggression, self-injurious behavior) in autistic children (Agt, Verhoeven, van den Brink, & de Koning, 2011; Chan, Williams, May, Wan, & Brignell, 2023; Coales et al., 2019; Feeney, Desha, Khan, & Ziviani, 2017). As such, there is reason to further investigate how dyadic coordination of vocalization and pauses may be associated with language abilities in autistic children.

Studies have reported a general reduction in synchronous verbal and nonverbal (e.g., eye-gaze, affect, joint attention) behaviors in autistic children interacting with their caregivers (Trevarthen & Daniel, 2005; Yirmiya et al., 2006; Zampella, Csumitta, Simon, & Bennetto, 2020). In one study that examined 11-month-old monozygotic twins, one of whom developed ASD at 2 years of age, exchanges of social behaviors (e.g., contingent vocalizations, eye-gaze, smiling, touch) between the caregiver and the infant later diagnosed with ASD had reduced instances of sequential, reciprocal, or cyclical social interactions compared to the typically developing infant (Trevarthen & Daniel, 2005). The infant who was later diagnosed with ASD displayed short bursts of reactions and engaged minimally in shared experiences, which in turn, led to specific interaction styles from their caregivers (e.g., relying on repeated physical stimulation to engage infant), which were distinct from the styles the caregivers used to engage with the typically developing infant. The authors highlighted that although both caregivers displayed affectionate behaviors and efforts to support each twin, reduced synchronous social interaction between the infant later diagnosed with ASD with her caregivers could undermine the growth of emergent behaviors critical for the development of early social and cognitive skills.

In the context of dyadic coordination of vocalization, day long recording devices worn by children in the home environment have revealed that the frequency of vocalizations between dyads with autistic children are reduced compared to dyads with typically developing children (Warlaumont et al., 2014). However, specific kinds of vocalization hypothesized to be optimal for speech development were present in both dyads with autistic and typically developing children; adults in both groups were more likely to contingently respond if the child's utterances contained speech-like-vocalizations. Moreover, the likelihood of the child's vocalization to contain speech-like material increased if the child's previous speech-like-vocalization received an adult response (i.e., social feedback loop; Warlaumont et al., 2014). Researchers speculated that the reduced frequency of vocalizations leading to weakened social feedback loops in dyads with autistic children may impact the quality or quantity of language input from the environment, thereby negatively impacting the growth of speech and language abilities in those children (Warlaumont et al., 2014). Taken together with additional empirical evidence of a bidirectional relationship between the caregivers' sensitivity (i.e., contingent response to child's behaviors) with the child's language abilities in infants with higher likelihood of ASD (Smith et al., 2023), there is a strong impetus to examine dyadic coordination of vocalization in relation to language abilities.

There is emerging evidence that certain aspects of the dyadic coordination of vocalization and pauses, such as vocalization that is simultaneous (i.e., overlapping) with a communication partner and between-speaker pauses (i.e., latency to respond), may be associated with later language delay in infants with autistic siblings. Northrup and Iverson (2015) studied low likelihood (i.e., no first or second-degree relatives diagnosed with ASD) and high likelihood (i.e., older sibling with confirmed ASD) infants, measuring the vocal coordination during interactions between these 9-month-old infants and their mothers. The authors found that the percentage of infant overlapping speech, which was defined as the frequency with which the infant interrupted the caregiver's vocalization divided by the sum of both caregiver and child vocalizations, was a significant predictor of language delay (i.e., standardized scores on language measures at or below the 10^{th} percentile at more than one time point between 18 to 36 months of age). That is, for every 1% increase in infant overlapping speech there was a 5.7% increase in the likelihood of the infant being language delayed at 2 and 3 years of age. Neither caregivers' overlapping speech (i.e., caregiver's vocalization interrupting and overlapping the infant's vocalization) nor infant ASD likelihood status predicted language delay. Furthermore, dyads in the high-likelihood but no language delay and low-likelihood groups demonstrated strong positive correlations between the caregivers' and infants' mean latency to respond, whereas language-delayed infants and their caregivers did not. Altogether, the dyadic coordination variables, a) child's overlapping speech, which was a significant predictor of language delay, and b) child's and caregiver's latency to respond (i.e., between speaker pauses) appear to be associated with language abilities in infants with high likelihood of ASD.

In the same study, a deeper examination of infants' overlapping speech and latency to respond in the high-likelihood group, both language delayed and no language delayed dyads, revealed that latency may be a unique factor that predicts global language scores (i.e., a composite score based on the MacArthur Bates CDI and MSEL; Fenson et al., 2007; Mullen, 1995). Northrup and Iverson (2015) calculated the mean duration of pauses attributed to either the caregivers or infants and derived an absolute value of the difference between the two means. In alignment with the assumption of the meta-theoretical construct "reciprocity" as proposed by Provenzi et al. (2018), the authors interpreted the difference score of 0 as close to "perfect coordination". In a linear regression model predicting the global language score from infants' overlapping speech, mean difference in latency to respond, and caregiver's education level, these three predictors accounted for 58% of the variance in language composite scores. Moreover, the difference in latency to respond was a significant and unique predictor of language composite score; for every 1 unit increase in the difference in latency to respond, language composite *z*-score would decrease by 0.53. As such, two particular variables in dyadic coordination (i.e., overlapping speech, difference in mean latency to respond) may be potential indexing factors to later global language scores in populations with family history of ASD.

Nonetheless, there are several aspects of dyadic coordination of vocalization and pauses in relation to language abilities that need further investigation and clarification. For example, unlike their study from 2015, Northrup and Iverson (2020) observed no difference in overlapping speech and latency between infants with autistic siblings and their caregivers when compared to dyads with typically developing children. It is unclear whether the lack of difference detected was due to particular sample characteristics or the study design. On the other hand, Feldstein et al. (1982) observed that adolescents with ASD who were engaged in conversations with their caregiver, or an examiner had longer between speaker pauses than adult dyads (i.e., caregiver - examiner dyads). While there was evidence that the adult dyads matched the lengths of between speaker pauses with one another, no such matching was detected in dyads with autistic adolescents. As such, there are some clues that between speaker pauses may present differently across speaker familiarity in older autistic individuals; however, the relationship between language abilities and between speaker pauses in this population remain unknown.

In sum, there is a paucity of literature that examines dyadic coordination of vocalization and pauses in autistic children despite evidence pointing towards its important relationship with language abilities. For instance, a clear distinction between an infant later diagnosed with ASD compared to her typically developing infant twin reveals how differences in verbal and nonverbal synchrony between infants and caregivers might be related to social and cognitive growth during the early stages of development (Trevarthen & Daniel, 2005). Further, findings of a weakened social feedback loop in dyads with autistic children suggest that reduced frequency of vocalization between the dyads may have a cascading impact on speech and language abilities

(Warlaumont et al., 2014). Finally, there is evidence that specific aspects of dyadic coordination of vocalization at 9 months of age, such as overlapping speech and difference in the mean latency to respond, may be critical factors predicting later language delays or global composite language scores (Northrup & Iverson, 2015). However, mixed findings and a lack of studies that examine the relationship between dyadic coordination of vocalization and pauses concurrently with language abilities in older autistic children warrants further investigation in this area.

Preliminary evidence

In a preliminary study, we examined overlapping speech in 24 children with ASD ages 3- to 11-years-old in relation to one specific domain of language – vocabulary. In addition to overlapping speech, we examined phonological memory, the ability to retain linguistic information. Although not a focus of the current study, phonological memory was included because it has been consistently related to limited vocabulary development in children with language impairment (Gray, 2004, 2006; Weismer et al., 2000), but research on phonological memory in children with ASD has been mixed (Hill, van Santen, Gorman, Langhorst, & Fombonne, 2015; Nadig & Mulligan, 2017; Norbury, Griffiths, & Nation, 2010). Both overlapping speech and phonological memory were extracted from a structured nonword repetition task administered in a laboratory setting, which included explicit directions on how to coordinate (e.g., "tell me the name of the monster after I say its name two times!"). We examined the child's percentage of overlapping speech in a highly structured task with clear instructions on the turn-taking aspect, as opposed to an unstructured natural environment, so that any differences in coordination of vocalization could be easily captured. We hypothesized that the dyadic coordination of vocalization in the highly structured task would predict receptive and expressive vocabulary scores measured using the Peabody Picture Vocabulary Test, Fourth

Edition (PPVT-4; L. M. Dunn & Dunn, 2007) and Expressive Vocabulary Test, Second Edition (EVT-2; Williams, 2007) .

Two separate linear regression analyses revealed that the set of predictors (nonverbal cognition, child's overlapping speech, and phonological memory) significantly accounted for variance in receptive and expressive vocabulary sizes. For every one standard deviation increase from the sample mean in overlapping speech, there were decreases of 8.42 points in PPVT-4 standard score and 15.39 points in EVT–2 standard score. As such, the negative relationship between the child's overlapping speech and language abilities (Northrup and Iverson, 2015), appears to persist in older autistic children engaged in structured tasks, into the school-age years.

There continues to be knowledge gaps in the presentation of dyadic coordination variables and its relationship to language abilities in school-aged autistic children. Our preliminary study examined the relationship between dyadic coordination variables and concurrent vocabulary size in school-aged autistic children. However, evidence of associations between dyadic coordination variables and composite language scores in 2 to 3-year-old children with high likelihood of ASD (Northrup & Iverson, 2015) warrants further investigation into the relationship between the dyadic coordination variables and composite language scores in older autistic children. In addition, there is evidence that dyadic synchrony presents differently in parent-child-play sessions compared to structured activities in young children attention deficit/hyperactivity disorders (ADHD) (Healey, Gopin, Grossman, Campbell, & Halperin, 2010). No study to this date has examined dyadic coordination of vocalization and pauses in parent-child-play sessions between school aged autistic children and their caregivers. As such, the current study aims to contribute to the knowledge gap by addressing the following primary research questions along with a series of exploratory analyses investigating dyadic coordination variables and language abilities in school aged autistic children in the context of natural environment.

Research Questions

1. Does dyadic coordination of vocalization and pauses with a caregiver, as indicated by overlapping speech and latency to respond, differ in autistic or typically developing children?
2. Does the dyadic coordination of vocalization and pauses predict concurrent language skills in school-age autistic or typically developing children?

Methods

The current study was a part of a larger study that examined word learning abilities in children with typical development or ASD (NIDCD R21DC017226; PI: Kover). The study implemented synchronous and asynchronous remote experimental paradigms for families across the U.S. In the synchronous portion of the study, families participated in two approximately 1.5 hour long online sessions with a 24-hour gap between the sessions. The synchronous sessions consisted of several language learning activities and assessments, including parent-child-play and language assessments via Zoom. For the asynchronous portion of the study, caregivers filled out the following questionnaires within two weeks before or after the synchronous meeting: a) Social Responsiveness Scale- Second Edition (SRS-2; Constantino & Gruber, 2012) and b) a family background questionnaire that captures race, ethnicity, caregiver education level, language environment, and socioeconomic status. REDcap (Research Electronic Data Capture; Harris et al. 2009; Harris et al., 2019), which is a secure, web-based platform, was utilized for parent-questionnaire completion.

Participants

Participants were school-age children with a known community diagnosis of ASD (*n* = 20; chronological ages 5;6 – 10;10 years) or typical development (*n* = 34; chronological ages 2;8 – 6;9 years) and their caregivers. Recruitment areas included the greater-Seattle with the aid of the UW Communication Studies Participant Pool, UW Autism Center, and community connections. Furthermore, recruitment extended nationally via social media platforms specifically used for research purposes in communication sciences and disorders and the ASD community (e.g., Facebook groups for families with children with ASD, PhD students in CSD, or SLPs in clinical research).

The following were the exclusion criteria for children with ASD or typical development: a) English was not the primary language spoken at home (i.e., the main language spoken at home must be English, but the child could be exposed to other languages through extended family or the school environment), b) the child cannot follow simple requests (e.g., sit in your chair, please), c) significant behavioral challenges that would preclude a positive experience in the study, d) has uncorrected physical or sensory limitations that would prevent the child from successfully participating in the activities of the research study. Children with typical development were not receiving any speech and language intervention, or special services.

There were no exclusionary criteria for caregivers. The caregivers' characteristics such as the household income and education level were extracted from self-reported family background questionnaires. See Table 1. The caregivers' gender was reported based on visual inspection of the parent-child-play sessions, thus, described as female or male presenting. The current study acknowledges the limitations behind caregiver gender assumptions and addresses this topic further in the Conclusions and Limitations section.

Table 1.
Demographic Information

	Autism Spectrum Disorder (*n* = 20)		Typical Development (*n* = 34)	
	Frequency	%	*Frequency*	%
Gender				
Male	12	60	16	47.1
Female	8	40	18	52.9
Child Race/Ethnicity				
White, non-Hispanic/Latino	9	45	26	76.5
White, Hispanic/Latino	2	10	1	2.9
Multiracial, non-Hispanic/Latino	3	15	5	14.6
Asian	1	5	0	0
Unreported	5	25	2	5.9
Caregiver gender in the PCP				
Male presenting	3	15	1	2.9
Female presenting	16	80	26	76.5
Missing data	1	5	7	20.6
Family income				
$25,000 - 50,000	1	5	0	0
$50,001-75,000	2	10	0	0
$75,001-100,000	2	10	1	2.9
$100,001-150,000	7	35	4	11.8
$150,001 - 250,000	0	0	14	41.2
More than $250,000	4	20	8	23.5
Unreported	4	20	7	20.6
Education				
Caregiver 1				
Terminal (Ph.D., M.D., MFA)	2	10	7	20.6
Graduate (M.A., M.S.)	5	25	10	29.4
College degree (B.A., B.S.)	5	25	11	32.4
Assoc/tech degree	3	15	0	0
Some graduate work	0	0	3	8.8
Some college/tech school	1	5	1	2.9
Unreported	4	20	2	5.9
Caregiver 2				
Terminal (Ph.D., M.D., MFA)	1	5	7	20.6
Graduate (M.A., M.S.)	6	30	8	23.5
College degree (B.A., B.S.)	4	20	10	29.4
Assoc/tech degree	3	15	4	11.8
Some graduate work	0	0	0	0
Some college/tech school	2	10	3	8.8
Unreported	4	20	2	5.9

Note: PCP = Parent-child-play session. The term "multiracial" includes individuals who reported a combination of White, Asian, Black, or African American, American Indian or Native Alaskan, or Native Hawaiian or Other Pacific Islander.

The Social Responsive Scale-2 (SRS-2; Constantino & Gruber, 2012) completed by caregivers, which provides interpretive scores applicable to specific age ranges, was used to describe the children's social communication skills in corroboration of the community ASD diagnosis (Frazier et al., 2023). The Preschool Form (ages 2.5 to 4.5 years) and the School-Age Form (ages 4 to 18 years), each standardized to a nationally representative sample within the specified age ranges, were used. Both the preschool and the school-age forms consist of 65 items that address social awareness, social cognition, social communication, social motivation, and restrictive repetitive behaviors appropriate to the specified age ranges. Caregivers rate each item on the questionnaire using a scale from 1 to 4 (1 = not true, 4 = almost always true). The total raw scores are converted to gender-normed T-scores; scores below 59 are considered within-normal-range, 60 – 75 are considered mild-to-moderately impaired, and above 75 are considered severely impaired. There was a total of five participants in the ASD group without SRS-2 total T-scores due to missing data (e.g., the parent did not complete the questionnaire). Three out of those five participants without SRS-2 data were included in the analysis for Research Question 1, and two were included in the analysis for Research Question 2. The limitations of this approach are addressed further in the Discussion. The SRS-2 T-scores are reported in Tables 1 and 2.

Table 2.
Direct Assessment Results and Parent Questionnaire Responses

	Autism Spectrum Disorder				*Typical Development*			
	n	*M*	*SD*	Range	*n*	*M*	*SD*	Range
Child's Age (months)	20	90.85	19.86	66 – 131	34	54.94	14.28	32 – 81
PPVT – 5 (GSV)	16	488.75	19.45	451 – 517	31	479.13	12.04	446 – 500
PPVT –5 (SS)	16	102.19	31.54	57 – 160	31	117.58	15.76	81 – 160
CELF-P3 (GSV)	11	559.15	28.28	505 – 591.67	25	549.37	22.87	493 – 583
CELF-P3 (SS)	3	30	8.54	22 – 39	24	38.13	4.73	23 – 47
SRS - 2	15	76.27	9.25	62 – 90	15	46.00	4.22	40 – 55

Note: PPVT-5 GSV = Peabody picture vocabulary test, 5th edition, growth scale values; PPVT-5 SS = Peabody picture vocabulary test, 5th edition, standard score; CELF- P3 GSV= Clinical Evaluation of Language Fundamentals

Preschool 3rd edition growth scale values; SS = standardized scores; SRS -2 = Social Responsive Scale-2 total T-scores.

Participant selection filters

In total, 54 dyads participated in the study (ASD *n* = 20; TD *n* = 34). Dyads with missing or incomplete parent-child-play sessions, receptive vocabulary scores, or composite language scores were excluded, using participant selection filters that maximized sample sizes and that were relevant to the primary analyses for Research Questions 1 and 2, and several exploratory analyses. Since the behavioral data extracted from parent-child-play sessions was required to answer both primary and exploratory research questions, dyads with missing parent-child-play sessions (ASD *n* = 1; TD *n* = 8) were excluded from all analyses.

Research Question 1 and establishing group equivalence

In Research Question 1, receptive vocabulary was measured using the Peabody Picture Vocabulary Test, Fifth Edition (PPVT-5; Dunn, 2019) and was required for matching the groups. There were 16 ASD dyads and 24 TD dyads with complete PPVT-5 growth scale values (GSV) and recorded parent-child-play sessions. To improve balance between the groups on the matching variable, one dyad from the TD group was excluded from analysis, with the child's PPVT-5 GSV score below 451, which was the lowest score in the ASD group. The final number of dyads for Research Question 1 was ASD = 16 and TD = 23. The difference in the PPVT-5 GSV of children in the ASD dyads (*M* = 488.75, *SD* = 19.45) and the TD dyads (*M* = 480.09, *SD* = 10.86) was non-significant, $t(21.53) = 1.63$, $p = .121$, Cohen's $d = 0.58$. Further, the difference in the PPVT-5 standard scores in the ASD dyads (*M* = 102.19, *SD* = 31.54) and the TD dyads (*M* = 117.70, *SD* = 15.16) was non-significant, $t(19.86) = -1.83$, $p = .083$, Cohen's $d = 0.67$.

Although the groups do not significantly differ in norm-referenced or absolute level of receptive vocabulary, the effect sizes indicate stronger absolute receptive vocabulary in the ASD group and stronger norm-referenced vocabulary in the TD children. This relates to the generally younger age of the TD children and vocabulary performance that lags behind chronological age expectations for ASD. This extent of overlap in distributions for vocabulary scores is not ideal (Kover & Atwood, 2013), but it does represent a conservative comparison between groups. In other words, any group differences in dyadic coordination could not be attributed to smaller absolute vocabulary size in the autistic children.

Research Question 2

In Research Question 2, composite language scores measured using the Clinical Evaluation of Language Fundamentals Preschool – 3rd edition (CELF- P3; Wiig, 2020) was required for data analysis. Dyads with children who lacked composite language scores (ASD n = 9; TD n = 9) were excluded specifically for Research Question 2, leaving 11 ASD dyads and 20 TD dyads with complete composite language scores (i.e., the dependent variable of the linear regression model) and valid parent-child-play sessions (i.e., the source of the independent variables). Research Question 2 did not involve group matching.

In a series of exploratory analyses, filters for Research Questions 1 or 2 were applied based on the nature of the analysis. When testing for group differences in the novel exploratory dyadic variable, difference in percent overlapping speech, the same filter from Research Question 1 (ASD n = 16; TD n = 23) was applied. When examining the association between the exploratory dyadic variable (independent variable) and composite language scores (dependent variable) in a multiple linear regression analysis, the filter from Research Question 2 was applied, as it was necessary for all dyads to have a composite language score.

Lastly, for the set of exploratory regression analyses using the child's receptive vocabulary as the outcome measure, first, the filter from Research Question 2 was applied to directly compare the results with the results from Research Question 2. Then, the sample was expanded to include all dyads with valid receptive vocabulary scores (ASD n = 16; TD n = 24) to examine generalizability of the regression model. Note, this expanded sample reintegrates the 1 TD dyad case that was originally excluded in the process of group matching for Research Question 1.

LANGUAGE MEASURES

An online version of the Peabody Picture Vocabulary Test, Fifth Edition (PPVT-5; Dunn, 2019) and the Clinical Evaluation of Language Fundamentals Preschool – 3rd edition (CELF- P3; Wiig, 2020) were administered using the Pearson Q-global website (Pearson Inc., 2023). The PPVT-5 quantifies receptive vocabulary skills by testing the participants' ability to listen to a verbal prompt (e.g., point to ball) to select the correct picture from an array of 4 full color drawings. The PPVT-5 yields standard scores and growth scale value. As described further below, growth scale values are akin to raw scores (i.e., measure absolute level of ability rather than relative performance in comparison to chronological age peers), but with preferred psychometric properties. The growth scale value was used to match the groups for Research Question 1 and as the outcome variable in the exploratory analysis.

The composite language score, used as the dependent variable in Research Question 2, was calculated using the core language subtests from the CELF-P3 (i.e., Sentence Structure, Word Structure, and Expressive Vocabulary). The total core language score is the sum of standardized scores of the core language subtests. However, since several participants (*n* = 8) were out of the testing age range and their scores could thus not be converted to standardized scores, the raw scores of the subtests were converted to growth scale values (GSV). The CELF-P3, rather than the CELF-5, had been selected for this study to ensure use of an omnibus language measure appropriate for children even with lower language levels. For the purpose of this study, the GSVs from the core language subtests were averaged and used as the composite language score in Research Question 2.

The growth scale values from both the CELF – P3 and the PPVT – 5 were used in all data analyses. Unlike the standard scores, the growth scale value represents the child's abilities based on an equal interval scale that is independent from the performances of the age-matched normative sample (Wiig, 2020). The growth scale value of a similar language assessment has been reported to have sound psychometric properties and sensitivity in detecting changes in language abilities of young autistic children (Kwok, Feiner, Grauzer, Kaat, & Roberts, 2022). Although the current study does not examine changes in language abilities, the growth scale values were selected to examine each participant's absolute language abilities, rather than their language abilities compared to a normative sample.

DYADIC COORDINATION OF VOCALIZATIONS AND PAUSES

Before the first synchronous session, caregivers were instructed to prepare toys that had similar characteristics to the following: stackable or build-able blocks, magnetic blocks or tiles, vehicles, small ball with ramp for rolling, pretend play toys (e.g., fruits, farm, doll house), animal figures or stuffed animals. Near the end of the first session, the examiner instructed the caregivers to use one or any of the toys they prepared to engage with their child in a natural way for 15 minutes. The examiner requested that the dyad engage in play near the electronic devices they were using for the session (e.g., laptops, computers, cellphone, tablets) to ensure good sound and visual quality of the parent-child-play recording. Once instructed, examiners engaged minimally with the dyads, only interrupting when they needed to be redirected to stay near their devices, reduce background noise, or reduce the presence of individuals not part of the study session.

Each parent-child-play session consisted of one adult and one child engaging with each other using toys of their choice in a natural environment. Most of the parent-child-play sessions were conducted in a home-like environment that did not appear to hinder natural interactions between the caregivers and their children. Two parent-child-play sessions from TD dyads were excluded due to poor audio quality as a result of internet connection issues or disruptive third parties. One parent-child-play session from the ASD group was excluded because the nature of the activity that the dyads engaged in (i.e., passive listening to book-reading) was different from the recommended list of toys in the parent instructions, and minimal vocal interaction was observed. One parent-child-play session in the ASD group included in the data analysis occurred in a quiet study room of a public library. Taking these exclusions into account yielded the 1 ASD and 8 TD dyad exclusions previously mentioned.

In similar studies that extracted behavioral data from caregiver-child interaction, 1 to 5-minute segments were sufficient in supplying the amount of data necessary for meaningful analysis (Gratier, 2003; Healey et al., 2010; Northrup & Iverson, 2020; Northrup & Iverson, 2015). Therefore, one approximately 5-minute long audio clip from each parent-child-play session, with minimal interruption from the environment (e.g., third party vocalizations, noises from toys) was extracted using the VLC media player (VideoLAN, n.d.) and conversational states (e.g., onset and offset of vocalization or pauses) were coded using the computer software Praat (Boersma & Weenink, 2019).

VOCALIZATIONS

The timing and types of vocalizations and pauses by the caregiver and child were coded using Praat (Boersma & Weenink, 2019). Voluntary verbal and nonverbal vocalizations made by the caregiver or child were coded as vocalizations (VOC) and attributed to the relevant speaker. Examples of verbal vocalizations include full or partial sentences, comments, praises, and fillers. Nonverbal vocalizations included grunting, screaming, laughing, and social gasping that was used to cue for attention or as a response. The end of the vocalization was marked if there was a pause within the speaker's utterance (intrapersonal gap) of longer than 300ms, or if there was a pause leading into the other speaker's vocalization (latency to respond). Other vocalizations, such as coughing, sneezing, non-social gasping, breathing, yawning, and blowing raspberries were not coded. The frequency and the mean duration of vocalizations can be considered conversational state variables, and were used to calculate one of the dyadic coordination variables of interest, percent overlapping speech. Refer to Table 3 for a complete list of all the conversational state and dyadic coordination variables.

Table 3.
Conversational State Variables and Dyadic Coordination Variables

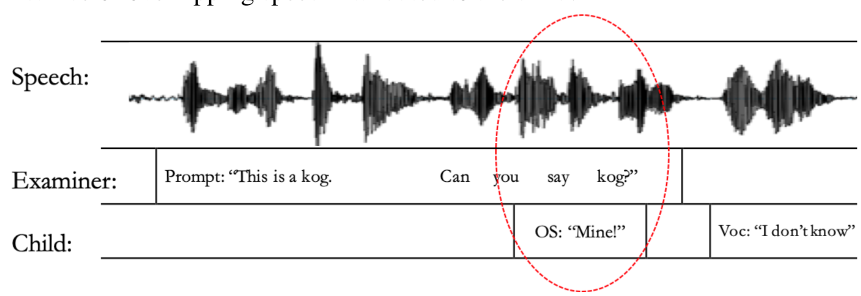
Variables	Definition
Conversational state variables	
Vocalization for caregiver or child	The total frequency and the mean duration (in seconds) of voluntary verbal and nonverbal vocalizations. Attributed to the speaker who makes the vocalization.
Intrapersonal pause for caregiver or child	The total frequency and the mean duration (s) of within-speaker pauses longer than 300ms. Attributed to the speaker who pauses between his/her vocalization.
Overlapping speech for caregiver or child	The total frequency and the mean duration (s) of overlapping speech attributed to the person who interrupted, creating overlapping vocalizations with the other speaker.
Latency to respond for caregiver or child	The total frequency and the mean duration (s) of between-speaker pauses attributed to the person who breaks the pause occurring at the end of the other speaker's vocalization.
Dyadic coordination variables	
Child's percent OS	$\frac{Child's\ freq.of\ OS}{Child's\ freq.of\ VOC + Caregiver's\ freq.of\ VOC}$
Caregiver's percent OS	$\frac{Caregiver's\ freq.of\ OS}{Child's\ freq.of\ VOC + Caregiver's\ freq.of\ VOC}$
Latency difference for the dyad	$\lvert\ Child's\ \mu\ dur.of\ latency - Caregiver's\ \mu\ dur.of\ latency\ \rvert$
Difference in percent OS for the dyad*	$\lvert\ Child's\ percent\ OS - Caregiver's\ percent\ OS\ \rvert$

Note: Freq. = frequency; $\mu\ dur.$ = mean duration; *difference in percent OS was a novel dyadic coordination variable used in one of the exploratory analyses.

PERCENT OVERLAPPING SPEECH

Overlapping speech (OS) occurs when both the child and the caregiver's vocalization overlap. For example, if the child interjected in the middle of his or her caregiver's vocalization, the moment in time when their vocalizations overlapped was coded as the child's OS. This coding method is similar to the one implemented by Northrup and Iverson (2015). See Figure 1 for a visual representation of overlapping speech. Both the frequency and the mean duration of OS were considered conversational state variables.

Figure 1. *Coding Scheme for Overlapping Speech*. The red circle indicates an instance of overlapping speech attributed to the child.



The percent overlapping speech in each dyad was calculated using the two conversational state variables: frequency of OS of the child or caregiver and the dyad's total number of vocalizations. The percentage of OS for each speaker (child; caregiver) was calculated by dividing the frequency of the speaker's overlapping speech by the total frequency of vocalizations within the dyad (child and caregiver combined). Percent overlapping speech, rather than frequency of overlapping speech, was utilized in analyses to account for variability across dyads in vocalization frequency, which could potentially impact the frequency of OS (Northrup & Iverson, 2015). The caregiver's and child's percent overlapping speech were two of the primary dyadic coordination variables of interest analyzed in Research Question 1 and Research Question 2. A novel dyadic coordination variable that examined the difference between the child's and caregiver's percent overlapping speech was used in one of the exploratory analyses. See Exploratory Analysis Results for more detail.

DIFFERENCE IN MEAN DURATION OF LATENCY TO RESPOND

Latency to respond (L), is the period of silence (i.e., pause duration) between when one speaker ends a vocalization, and the other speaker begins their vocalization. Similar to overlapping speech, latency was coded such that the pause duration was attributed to the speaker who broke the pause. For example, when the child ends their vocalization and the caregiver begins his vocalization afterwards, the period of silence between the two speaker is labeled "L" and attributed to the caregiver (i.e., the person who broke the pause by responding to the child's previous vocalization). No maximum boundaries for the duration of latency were set, but minimum length of latency was set at 10ms because any gap between two vocalizations shorter than that was not considered perceivable. As such, latency to respond was not coded between vocalizations that occurred back-to-back with less than a 10ms gap between speakers.

In addition to latency, pauses within speakers' vocalizations were coded as intrapersonal gaps (IAG). Intrapersonal gaps were defined as a pause longer than 300ms within the same speaker (Gratier, 2003; Northrup & Iverson, 2015). Both the frequency and the mean duration of latency to respond and intrapersonal gaps were considered conversational state variables.

The final dyadic coordination variable of interest, called *latency difference* from here on, is the absolute value of the difference between the caregivers' and child's mean duration of latency to respond. There was strong evidence that latency difference between infant and caregivers predicted global language scores (i.e., Mullen Scales of Early Learning (MSEL); Mullen, 1995) in infants with autistic siblings (Northrup & Iverson, 2015). Using a similar definition as the latency difference used in Northrup and Iverson (2015), the difference score of zero was interpreted as "perfect coordination" (Beebe et al., 1988; Northrup & Iverson, 2015), whereas a larger latency difference between the caregiver and the child may be indicative of asynchronous interaction. Refer to Table 3 for an organization of all the conversational state variables and dyadic coordination variables of interest.

In sum, the total frequency, and the mean duration of conversational state variables (vocalizations, intrapersonal pauses, overlapping speech, latency to respond) attributed to either the caregiver or the child were extracted from the raw data. Using these conversational state variables, 3 dyadic coordination variables (child's and caregiver's percent overlapping speech, latency difference) were calculated for each dyad. One additional novel dyadic coordination variable was generated to further explore overlapping speech between dyads.

CODING PROCEDURES AND TRAINING

One of the two trained and independent coders annotated each of the 5–minute audio clips using the computer software Praat (Boersma & Weenink, 2019). The annotation procedure consisted of determining the start and end time of vocalization behaviors (i.e., segmentation) and labeling each segment as specific vocalization behaviors according to their conversational state (e.g., VOC, L, OS).

A training process on a subset of files established that the coders were sufficiently prepared for independent coding. The training process involved becoming familiar with the code definitions, independently annotating audio clips, comparing annotations between the lead researcher and the undergraduate research assistant, resolving discrepancies, and revising codes as necessary until reaching approximately 90% accuracy in segmenting and labeling vocalizations and pauses. The interrater reliability of labeling on three training files were established using Cohen's Kappa, ϰ = .934 (95% CI), p < .001, which indicates substantial agreement. In general, approximately 85% of the start and end segmentations were at or less than 100ms different.

INTER-RATER RELIABILITY

Once the coding scheme was finalized and independent coding commenced, roughly 20% of the audio files (ASD n = 5; TD n = 7) were independently annotated by the reliability coder to assess agreement with the primary coder. Cohen's kappa was used to assess the reliability of categorical behavioral data (i.e., vocalization and pause labels). Cumulative distribution functions were used to assess the reliability of continuous data (i.e., start and end times of vocalizations and pauses). In both the TD (ϰ = .0.763 [95% CI], $p < .001$) and ASD group (ϰ = .0.791 [95% CI], $p < .001$), agreement on the categorical labels for vocalizations and pauses in the child and caregiver tiers was substantial, as it was between the range of 0.61 and 0.80 (Cohen, 1960).

To test the reliability in the segmentation of vocalization and pause behaviors, all behavior segments from the primary coder were manually aligned to the behavior segments of the reliability coder across time. Roughly 14% of segmentations in the TD group and 12% in the ASD group were misaligned (e.g., coder A did not detect the same behaviors as coder B; for example, one coder divided one long vocalization into two vocalization segments). For the segments that were aligned, the absolute difference between the primary and the reliability coder's start or end timepoints were calculated in seconds and milliseconds. The distribution of differences in start or end timepoints was examined using cumulative distribution functions. A cumulative distribution function has previously been used in research on a different vocal behavior (voice onset time), but nonetheless compared the beginning and end timepoint of two independent coders using Praat (Ryant, Yuan, & Liberman, 2013). As such, the same method was implemented in the current study to quantify the level of agreement, as well as display a visual representation of the alignment of the codes.

Among the segmentations that were aligned in the TD group, the percent occurrence of the difference between coders being less than or equal to 100ms was 87.2% for the start times and 82.52% for the end times. In other words, among the segments that were aligned between the two coders, approximately 85% of start or end timepoints of the segments were within 100ms of each other. Similarly, in the ASD group, 86.34% of the start times and 83% of the end times were within 100ms different (See Figure 2).

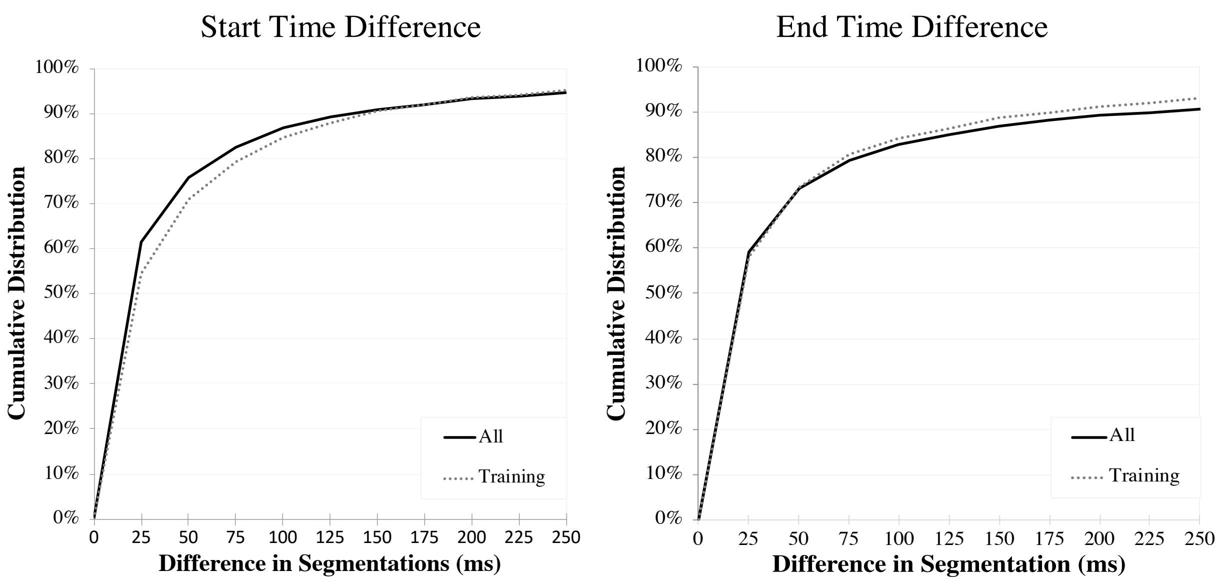


Figure 2. *Cumulative Distribution Function of Start and End Time Differences*. This figure shows the cumulative distribution of differences in the start or end time segmentations by two independent coders with training data (n = 3) and all reliability coded data (n = 12)

Although the percent occurrence for agreement is lower for shorter time windows (e.g., approximately 75 percent for 50ms for start timepoints), 100ms was selected as the most appropriate in relation to the observed data. To contextualize whether 100ms is an appropriately sized window for acceptability of timing alignment of vocalization segmentation, a cumulative distribution analysis of the duration of all behavioral codes across 93% of the participant dyads (*n* =42) was completed. That analysis shows that 95% of the vocalizations were at or below 3 seconds long and less than 0.5% were shorter than 100ms. Likewise, for the duration of overlapping speech and latency, approximately 99% of overlapping speech intervals and 88% of latency intervals were at or below 2 seconds long and less than 10% of both types were shorter than 100ms. See Figures 3 and 4 to examine the duration of all the behavioral codes in seconds and in milliseconds, respectively. As such, it is unlikely that differences less than 100ms long between the two coders would significantly impact the accuracy of the behavioral codes.

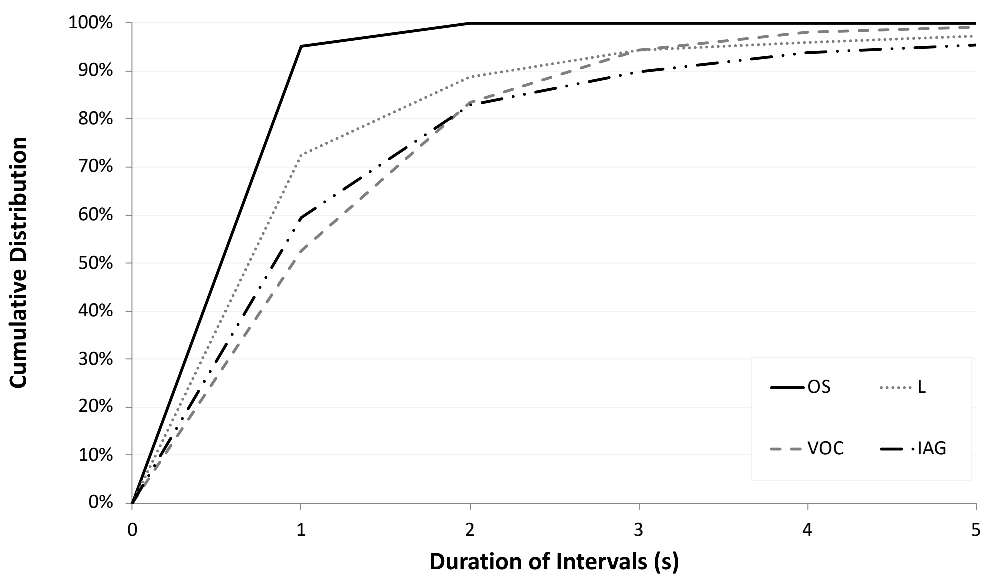


Figure 3. *Cumulative Distribution Function of all Behavioral Codes in Seconds.* This figure shows the cumulative distribution of overlapping speech (OS), latency (L), vocalizations (VOC), and intrapersonal pauses (IAG) from 0 to 5s. $N = 42$.

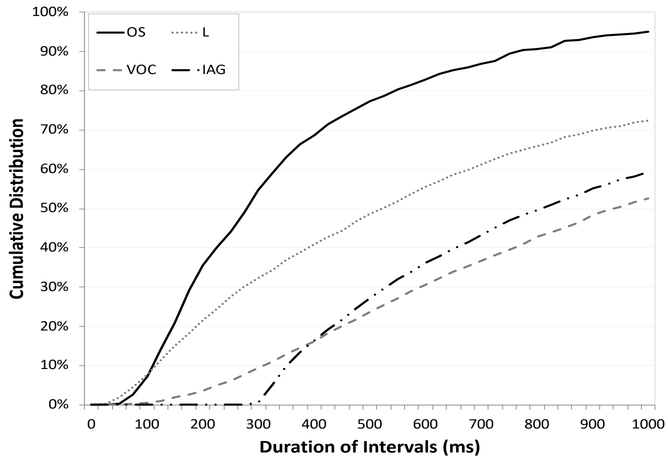


Figure 4. *Cumulative Distribution Function of all Behavioral Codes in Milliseconds*. This figure shows the cumulative distribution of overlapping speech (OS), latency (L), vocalizations (VOC), and intrapersonal pauses (IAG) from 0 to 1000ms.

ANALYSIS PLAN

Regarding Research Question 1, before testing for group differences in dyadic coordination and conversational state variables, variance and normality of the groups matched on receptive vocabulary score were tested using the Levene's Test of Equality of Variances and the Shapiro-Wilk test. For the dyadic coordination variables (e.g., caregiver and child's percent overlapping speech, latency difference), the Levene's Test of Equality results showed that equal variances could be assumed (all $p > .05$). However, the Shapiro-Wilk test showed that all except the caregiver's percent of overlapping speech ($W = 0.95$, $p = 0.077$), had a significant p – value (ps $< .05$) indicative of deviation from normal distribution. As such, the Mann-Whitney U-test was used to test for group differences in child's percent overlapping speech and latency difference, whereas the independent sample *t*-test was used for the caregiver's percent overlapping speech. The Mann-Whitney U test is a nonparametric test used as an alternative to the independent samples t – test when assumptions of normality are not appropriate (McKnight & Najab, 2010).

The same procedure was applied to test for equal variances and normality in conversational state variables: frequency and mean duration of vocalization, latency between speakers, intrapersonal gaps, and overlapping speech. All frequency and mean duration variables had equal variances (all $p > .05$). As for normality, the caregiver's frequency of overlapping speech ($W = 0.98, p < .010$), child's frequency of overlapping speech ($W = 0.83, p < .001$), caregiver's mean duration of latency ($W = 0.90, p = 0.002$), mean duration of overlapping speech ($W = 0.89, p = .001$), and mean duration of intrapersonal gap ($W = 0.94, p = 0.035$), child's mean duration of vocalization($W = 0.92, p = .008$), latency ($W = 0.89, p = .001$), mean duration of overlapping speech ($W = 0.94, p = 0.03$), and mean duration of intrapersonal gap ($W = 0.84, p < .001$), deviated from normal distribution. As such, the Mann-Whitney U test was used to test for group differences in the aforementioned variables, and the independent sample t-test was for all normally distributed variables (i.e., caregiver's frequency of vocalization, frequency of latency, frequency of intrapersonal gap and the mean duration of vocalization; child's frequency of vocalization, frequency of latency, and frequency of intrapersonal gap).

Regarding Research Question 2, to test whether dyadic coordination of vocalizations and pauses predicted concurrent language abilities, the dyadic coordination variables (caregiver's/child's percent overlapping speech and latency difference) were considered as potential predictors of the child's composite language score. Groups were effect coded (ASD = -1, TD = +1) so that the regression coefficient for group would be the average composite language score across both groups. Furthermore, the primary variables were standardized into z-scores for better interpretation of the regression coefficients. The Pearson r was calculated to determine the level of correlation between the dependent variable (composite language score) with the independent variables (caregiver's/ child's percent overlapping speech, latency difference, group). Due to limited predictor-to-sample ratio of 15:1 (Tabachnick & Fidell, 2012), the final regression model only included 3 independent variables: Group, child's percent overlapping speech, and latency difference. The caregiver's percent overlapping speech was excluded from the final model as it was not significantly correlated with the dependent variable in the current study and was also nonsignificant in a similar analysis conducted by Northrup and Iverson (2015). Although the child's percent of overlapping speech was also not significantly correlated to the dependent variable in the current study, it was included in the final model as there was strong theoretical support for the variable based on previous research (e.g., it was a significant correlate of vocabulary size in a preliminary study; it approached significance in a similar analysis conducted by Northrup and Iverson (2015).

$$Y_{\text{CELF Growth Score}} = b_0 + b_1 * \text{Group} + b_2 * \text{Child's percent OS} + b_3 * \text{Latency Difference} + e$$

EXPLORATORY ANALYSIS PLAN

DIFFERENCE IN PERCENT OVERLAPPING SPEECH

In addition to the dyadic coordination variables of interest, the absolute difference in the child's and caregiver's percent overlapping speech, a novel exploratory dyadic variable, was calculated for each dyad. Similar to latency difference, which was proposed by Northrup and Iverson (2015), a value of 0 was interpreted as "perfect coordination" (e.g., both the child and caregiver talk over each other frequently), whereas values greater than 0 would be indicative of asynchronous interaction (e.g., one person overlapped the other frequently and the other person never overlapped them). The Mann-Whitney U test was used to detect any group differences. Further, in an exploratory analysis, the difference in percent overlapping speech was included in place of child's percent overlapping speech in primary Research Question 2, to test for an association with composite language abilities. The following regression model, which included two dyadic variables, was tested in dyads with valid composite language scores and parent-child-play sessions.

$$Y_{\text{CELF Growth Score}} = b_0 + b_1 * \text{Group} + b_2 * \text{Difference in percent OS} + b_3 * \text{Latency Difference} + e$$

RECEPTIVE VOCABULARY AS THE OUTCOME

Parallel to the results from a preliminary study on overlapping speech and concurrent vocabulary size (under review; Choi & Kover, 2021), we explored how the dyadic coordination variables from the current study were associated with concurrent receptive vocabulary size. In the preliminary study, the child's overlapping speech was significantly and negatively associated with receptive and expressive vocabulary score, while keeping nonverbal cognition and phonological memory constant. The potential association between dyadic coordination and vocabulary in particular is of theoretical interest because overlapping speech may potentially impact the quality and quantity of language learning opportunities (e.g., influence the ability to learn words from that input), thus presumably limiting the child's vocabulary size. In the current study, the following regression model was first analyzed, in the same sample as Research Question 2, as an exploratory analysis. Then, to increase sample size and test for generalizability, participants with existing receptive vocabulary scores, but without composite language scores, were included in the analysis.

$$Y_{\text{PPVT Growth Score}} = b_0 + b_1 * \text{Group} + b_2 * \text{Child's percent OS} + b_3 * \text{Latency Difference} + e$$

RESULTS

DESCRIPTIVE ANALYSIS: DYADIC COORDINATION BETWEEN CHILD AND CAREGIVER

A comprehensive conversational state variables used in the current study have not been previously utilized in dyads with school-age children and their caregivers engaged in parent-child-play sessions. As such, a descriptive analysis of the conversational state variables (frequency and latency of overlapping speech and latency) that pertained to the dyadic coordination variables, and the child and caregiver's percent overlapping speech was used to provide context for the results of the primary analyses. See Tables 4 and 5.

Table 4.
Descriptive Results of Conversational State Variables

	Autism Spectrum Disorder				*Typical Development*			
Caregivers	*n*	*M*	*SD*	Range	*n*	*M*	*SD*	Range
Frequency								
Vocalization	19	72.11	19.37	40 – 108	26	68.46	24.19	9 – 107
Latency	19	27.42	10.63	2 – 45	26	33.08	12.36	7 – 66
Overlapping Speech	19	10.47	7.43	2 – 29	26	10.88	6.99	0 – 27
Intrapersonal Gap	19	35	18.43	5 – 68	26	23.69	14.21	1 – 55
Duration (s)								
Vocalization	19	1.14	0.21	0.92 – 1.55	26	1.26	0.34	0.57 – 2.08
Latency	19	1.3	0.89	0.39 – 3.73	26	0.79	0.39	0.24 – 1.81
Overlapping Speech	19	0.43	0.17	0.19 – 0.9	26	0.35	0.14	0 – 0.81
Intrapersonal Gap	19	1.91	0.82	0.64 – 3.36	26	1.47	0.61	0.71 – 2.84
Children	*n*	*M*	*SD*	Range	*n*	*M*	*SD*	Range
Frequency								
Vocalization	19	60.68	25.5	7 – 113	26	79.96	17.67	43 – 115
Latency	19	30.53	12.47	2 – 55	26	37.65	12.1	7 – 67
Overlapping Speech	19	9.47	5.29	1 – 21	26	11.46	9.63	0 – 44
Intrapersonal Gap	19	23.95	15.39	4 – 54	26	34.77	15.69	11 – 72
Duration (s)								
Vocalization	19	1.14	0.44	0.48 – 2.19	26	1.19	0.24	0.89 – 1.83
Latency	19	1.6	1.57	0.39 – 7.74	26	1.17	0.65	0.49 – 1.17
Overlapping Speech	19	0.34	0.13	0.18 – 0.65	26	0.29	0.13	0 – 0.61

Intrapersonal Gap	19	1.39	0.57	0.67 – 2.65	26	1.22	0.60	0.56 – 3.64

Note: All participants with parent-child-play sessions $N = 45$.

Table 5.
Descriptive Results of Dyadic Coordination Variables

	Autism Spectrum Disorder				*Typical Development*			
	n	*M*	*SD*	Range	*n*	*M*	*SD*	Range
Caregiver's OS (%)	19	7.30	4.32	2.41 – 18.12	26	7.26	4.5	0 – 18.75
Child's OS (%)	19	6.90	3.37	1.27 – 15.22	26	7.33	5.66	0 – 22.11
Latency Difference (s)	19	0.63	0.87	0.02 – 4.01	26	0.48	0.48	0 – 1.89
Difference OS (%)	19	2.7	2.4	0 – 10.11	26	3.65	3.32	0 – 10.72

Note: OS = overlapping speech; Difference in OS was examined as an exploratory analysis

All participants with parent-child-play sessions, even those with missing receptive vocabulary scores or composite language scores (ASD $n = 19$; TD $n = 26$), were included in these descriptive analyses. In both groups, two conversational state variables (frequency of overlapping speech and latency) were positively and significantly correlated between the caregivers and children. Frequency of overlapping speech of children significantly and positively correlated with the caregiver's frequency of overlapping speech in the ASD dyads, $r(17) = .64$, $p = 0.003$, and the TD dyads $r(24) = .64.$, $p < .001$. The frequency of latency, attributed to the person who interrupts the between-speaker pause, was also significantly and positively related between the caregiver and child in the ASD dyads, $r(17) = .86$, $p < .001$, and the TD dyads $r(24) = .84.$, $p < .001$. See Figure 5.

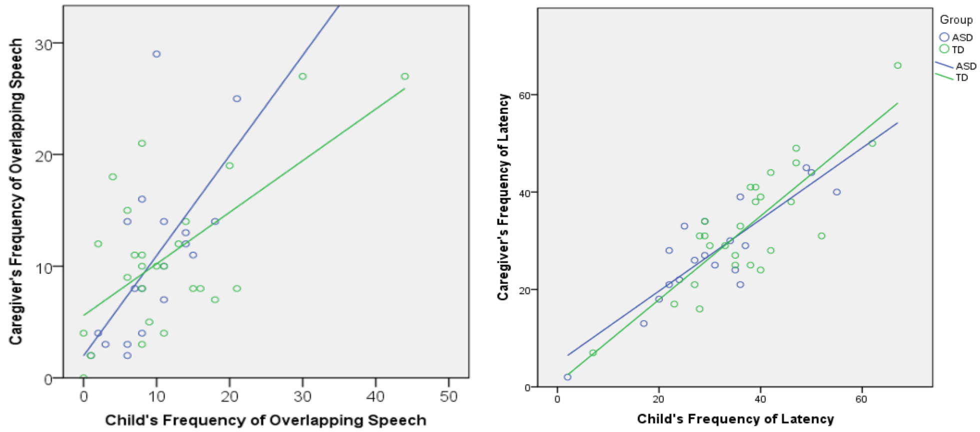


Figure 5. *The Child's and Caregiver's Frequency of Overlapping Speech and Latency*. This figure shows the correlation between the child's and caregiver's frequency of overlapping speech and latency from all parent-child-play sessions (ASD n = 19; TD n = 26).

Similar trends were observed in the duration of latency and overlapping speech between caregivers and their children. The duration of overlapping speech was positively correlated between caregivers and their children in the TD dyads, $r(24) = .77$, $p < .001$. Caregivers and children in the ASD dyads also displayed positive relationship in the duration of overlapping speech, but the relationship did not reach significance $r(17) = .35$, $p = .147$. The caregivers' duration of latencies was also positively correlated with the children's duration of latencies in both the ASD, $r(17) = .783$, $p < .001$ and TD dyads, $r(24) = .46$, $p = .018$. See Figure 6.

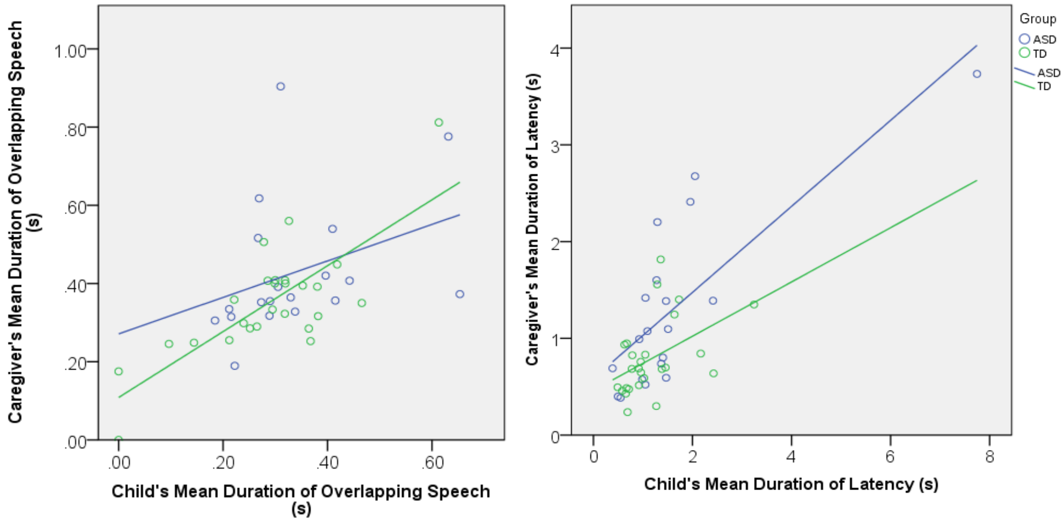


Figure 6. *The Child's and Caregiver's Duration of Overlapping Speech and Latency*. This figure shows the correlation between child's and caregiver's mean duration of overlapping speech and latency from all parent-child-play sessions (ASD n = 19; TD n = 26).

Lastly, the percent of overlapping speech, a dyadic coordination variable, was also significantly and positively correlated between children and their caregivers in the $r(17) = .57$, $p = .01$ and TD group, $r(24) = .54$, $p = .005$. The positive relationship between the caregiver's and child's frequency and mean duration of latency and overlapping speech suggests an alignment of vocalization behaviors in caregiver-child interactions in dyads, regardless of diagnostic status. See Figure 7.

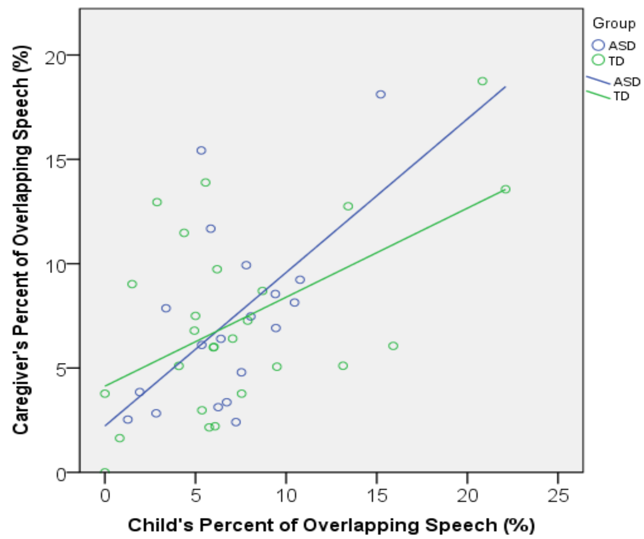


Figure 7. *The Child's and Caregiver's Percent Overlapping Speech*. This figure shows the correlation between child's and caregiver's percent overlapping speech from all parent-child-play sessions (ASD *n* = 19; TD *n* = 26).

Primary analysis

Research Question 1

This research question examined group differences in the dyadic coordination variables (caregiver's/child's percent overlapping speech and latency difference), and additional conversational state variables (frequency and mean duration of vocalizations and pauses). For dyads matched on the child's receptive vocabulary score (ASD = 16, TD = 23), there were no significant differences in any of the dyadic coordination variables. There were no differences detected in the caregiver's percent overlapping speech($t(37) = -0.32$, $p = .751$), child's percent overlapping speech ($U = 165$, $z = -0.54$, $p = .601$), and latency difference ($U = 167$, $z = -0.49$, $p = .641$) between the ASD and TD dyads.

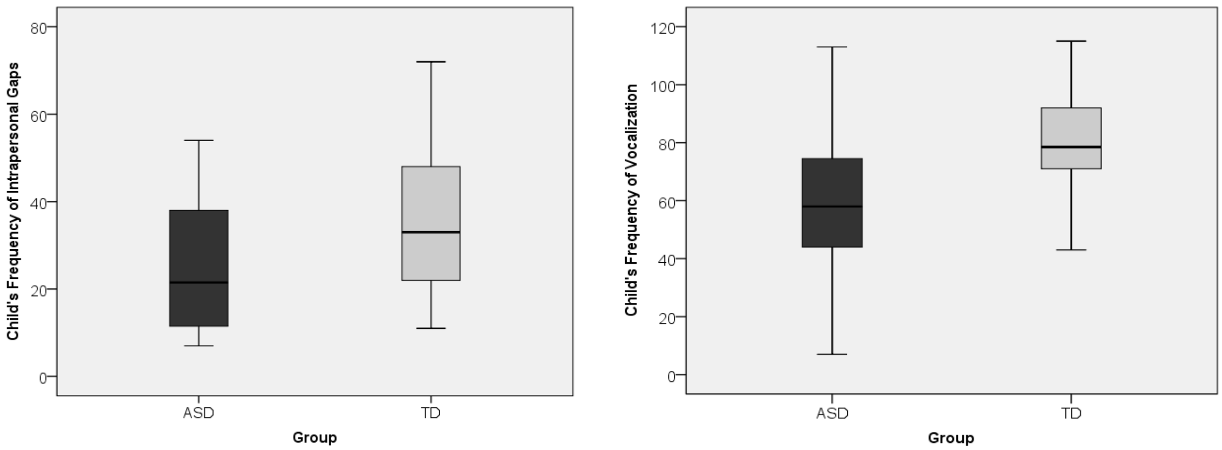


Figure 8. *Child's Frequency of Intrapersonal Gaps and Child's Frequency of Vocalization by Group*. This figure shows the significant group differences in the child's frequency of intrapersonal gaps and vocalization.

When the caregiver's and child's conversational state variables (frequency and mean duration of vocalization and pauses) were compared between the groups, there was a significant difference in the frequency of child's vocalization, $t(37) = 2.52$, $p = .016$, $d = 0.82$. The mean frequency of vocalization of autistic children was lower ($M = 64.06$; SD = 21.62) than that of typically developing children ($M = 79.57$; SD = 16.85). Also, the child's frequency of intrapersonal gaps (i.e., pauses within the same speaker) in the ASD group ($M = 25.13$, $SD = 15.34$) was significantly less from the TD group ($M = 35.74$, $SD = 15.94$), $t(37) = 2.08$, $p = .045$, $d = 0.68$. See Figure 8. No significant differences were observed in any of the caregiver's conversational state variables, but note that caregiver's intrapersonal gaps between groups had an effect size of $d = .65$, with the direction of that effect opposite to children's intrapersonal gaps. One possibility to explain those differences in direction of effect is that caregivers of children with ASD are waiting more frequently for their child to respond, but when the child doesn't respond, the caregiver speaks again, making that pause an intrapersonal gap rather than latency to respond; however, this has yet to be confirmed in the data. See Tables 6 and 7.

Table 6.
Independent t-test Results for Dyadic Coordination and Conversational State Variables

	ASD M (SD)	*TD* M (SD)	*t*	*p*
Dyadic coordination				
Caregiver's Percent OS	7.81 (4.40)	7.33 (4.72)	-0.32	.751
Conversational state				
Caregivers				
Frequency				
Vocalization	71.94 (20.93)	66.04 (23.61)	-0.80	.427
Latency	28.88 (8.69)	31.83 (10.95)	0.90	.372
Intrapersonal Gap	32.69 (18.07)	22.61 (13.48)	-1.99	.053
Duration (s)				
Vocalization	1.14 (0.20)	1.28 (0.36)	1.35	.185
Children				
Frequency				
Vocalization	64.06 (21.62)	79.57 (16.85)	2.52	.016
Latency	32.63 (10.87)	36.22 (11.16)	0.99	.324
Intrapersonal Gap	25.13 (15.34)	35.74 (15.94)	2.08	.045

Note: Groups matched on receptive vocabulary score (ASD n = 16; TD n = 23); OS = Overlapping speech

In sum, no group differences were detected in any of the dyadic coordination variables (i.e., caregiver's and child's percent overlapping speech, latency difference). Further, no group differences were detected in the frequency or mean durations of caregiver's vocalization and pauses. While there were some group differences in the child's frequency of vocalization and intrapersonal pauses, the majority of conversational state variables were not significantly different from each other.

Table 7. *Results from the Mann-Whitney U for Dyadic Coordination and Conversational State Variables*

	ASD Mean Rank	*TD* Mean Rank	*z*	*p*
Dyadic coordination				
Child's Percent OS	21.19	19.17	-0.54	.601
Latency Difference	21.06	19.26	-0.485	.641
Difference in POS	19.12	20.61	0.40	.703
Conversational state				
Caregivers				
Frequency				
Overlapping Speech	20.38	19.74	-0.17	.877
Duration (s)				
Latency	23.06	17.87	-1.40	.168
Overlapping Speech	23.38	17.65	-1.54	.128
Intrapersonal Gap	23.50	17.57	-1.60	.114
Children				
Frequency				
Overlapping Speech	19.78	20.15	0.10	.921
Duration				
Vocalization	18	21.39	0.91	.373
Latency	22.81	18.04	-1.29	.207
Overlapping Speech	20	20	0	1
Intrapersonal Gap	22	18.61	-0.91	.373

Note: Groups matched on receptive vocabulary score (ASD n = 16; TD n = 23); POS = Percent overlapping speech

Research Question 2

This research question addressed the association between dyadic coordination and language ability among participants. The regression analysis was preceded with bivariate correlations. See Table 7. For dyads with complete parent-child-play sessions and composite language scores, a (ASD n = 11 and TD = 20), only latency difference was positively and significantly correlated with composite language score, r (29) = .39, p = .016. Contrary to our hypothesis, children with higher composite language scores demonstrated greater difference in the mean duration of latency between the dyads. Neither group, $r(29) = -0.17$, $p = .179$, nor child's percent overlap, $r(29) = -0.15$ $p = .211$, correlated with the child's composite language score. Results did not appear to be impacted by multicollinearity (i.e., intercorrelations) between the predictors, as none of the correlation values between the predictors exceeded ± 0.90 (Tabachnick & Fidell, 2012).

Table 7.

Pearson r Correlation Between Composite Language Score and Primary Variables

Measure	1.	2.	3.	4.
Outcome				
1. CELF - P3 GSV	--			
Predictors				
2. Group	-0.17	--		
3. Child's percent OS	-0.17	-0.08	--	
4. Latency Difference	.39*	.08	-0.32	--

Note. OS = overlapping speech

* $p < .05$

The set of predictors (group, child's percent overlapping speech, and latency difference) did not account for significant variance in composite language score, ($R^2 = 0.19$, $F(3,27) = 2.13$, $p = 0.120$, $R^2_{adjusted} = 0.10$). See Table 8.

Table 8.
Model Fit and Beta Coefficients of the Linear Regression Analysis in Research Question 2

	R^2_{total}	$R^2_{adjusted}$	$F(3,27)$	p
CELF-P3 GSV	0.19	0.10	2.13	.120
	b	*(SE)*	*t*	*p*
Intercept	555.01	(4.44)	124.88	< .001
Group	-5.21	(4.46)	-1.17	.253
Child's Percent OS	-1.00	(4.58)	-0.22	.829
Latency Difference	9.71	(4.57)	2.12	.043

Note. *N*=31. CELF -P3 GSV = growth scale value; Group was effect coded (ASD = -1, TD = +1) and child's percent overlapping speech (OS) and latency difference were standardized into *z* -scores.

EXPLORATORY ANALYSIS

EXPLORATORY DYADIC VARIABLE; DIFFERENCE IN PERCENT OVERLAPPING SPEECH

In relation to Research Question 1, an exploratory dyadic coordination variable was tested for a group difference. Between groups matched on receptive vocabulary score (ASD n = 16; TD n = 23), there was no difference detected in the absolute difference of the child's and the caregiver's percent overlapping speech (U = 198, z = 0.40 p = 0.703). This suggests that the caregivers and children from the ASD group were engaging in similar alignment in proportions of overlapping speeches as dyads in the TD group. The same dyadic variable was also tested for its association with concurrent composite language scores.

The child's percent overlapping speech was replaced with the difference in percent overlapping speech, while keeping the other predictors (group, latency difference) and the outcome variable (CELF-P3 GSV) same in the regression model. A bivariate correlation between the predictors and the outcome revealed that only latency difference was positively and significantly correlated with composite language score, *r* (29) = .39, *p* = .016. Group membership, *r* (29) = -0.17, *p* = .179, and difference in percent overlapping speech, *r* (29) = .03, *p* = .431, were not significantly correlated with concurrent composite language scores.

Among dyads with valid composite language scores (ASD *n* = 11; TD *n* = 20), the set of predictors (group, difference in percent overlapping speech, and latency difference) did not account for significant variance in the child's composite language score, (R^2 = 0.20, *F* (3,27) = 2.19, *p* = .112, $R^2_{adjusted}$ = 0.106). As such, the exploratory dyadic variable does not appear to be a significant factor contributing to concurrent composite language abilities in older children in either the ASD or TD groups.

Exploratory regression analyses; receptive vocabulary as the outcome

In relation to Research Question 2 and given the theoretical and empirical evidence from a preliminary study and current analysis, we explored the association between the child's percent overlapping speech, latency difference, and vocabulary using PPVT-5 growth scale values as the outcome measure. These analyses were completed with two different selection criteria of participants: a stricter selection based on complete language scores and a broader selection with valid language scores only for the exploratory outcome (receptive vocabulary).

Receptive vocabulary as the outcome, tested on dyads with valid composite languages scores

The first exploratory regression model with receptive vocabulary as the outcome was tested in the same sample as Research Question 2 (i.e., dyads with complete parent-child-play sessions and CELF-P3 growth scale values; ASD = 11, TD = 20). The PPVT-5 Growth Scale Value (GSV) was used as the outcome. A bivariate correlation analysis between the outcome and the predictors revealed that difference in latency was positively and significantly related to receptive vocabulary score r (29) = .36, p = .024. Child's percent overlapping speech r (29) = .141, p = .225 was not correlated with receptive vocabulary.

Group membership was also significantly and negatively correlated with vocabulary score r (29) = -0.43, p = .008. In other words, the ASD group (effect coded as -1) had higher PPVT-5 GSV than the TD (+1) group. Note, group membership was not a significant factor in Research Question 2, when composite language scores were used as the outcome variable. In other words, in the sample of dyads for the current analysis and Research Question 2, children from the ASD group appear to have a significantly higher receptive vocabulary scores than children in the TD group, but that difference is not observed in their composite language scores. To better understand the effects of group membership in the overall model, a sequential multiple linear regression analysis was selected with group entered in the first block, and the dyadic coordination variables entered in the second block. See Table 9.

Table 9.
Multiple Linear Regression with Sequential Predictor Entry for Exploratory Analysis 1; receptive vocabulary as the outcome variable

	Block 1					Block 2				
	R^2_{change}	R^2_{total}	R^2_{adj}	b	sr^2	R^2_{change}	R^2_{total}	R^2_{adj}	b	sr^2
Model Fit	0.18*	0.18*	0.15			0.21*	0.40**	0.33		
Coefficients										
Intercept				487.19***					487.19***	
Group				-5.68*	0.18				-5.88**	0.19
Child's Percent OS									3.35	0.06
Latency Difference									6.17**	0.20

Note. N=31. Block 1 F-change test df = 1,29; Block 2 df = 3, 27; OS = Overlapping speech; Outcome measure was the PPVT – 5 growth scale value.
* p < .05, ** p < .01, *** p < .001.

Multiple linear regression with sequential predictor entry was used to test for incremental variance in receptive vocabulary explained by each predictor (group, child's percent overlapping speech, latency difference. Results showed that group membership, which was entered first, accounted for significant variation in receptive vocabulary score (R^2 = 0.18, F(1,29) = 6.47, p = .017, $R_{Adjusted}$ = 0.15). Controlling for group membership, the main effects of child's percent overlapping speech and latency difference accounted for 20% of variance in receptive vocabulary, R^2_{change} = 0.21, F_{Change}(3,27) = 4.78, p = .017. Altogether, the set of predictors (group, child's percent overlapping speech and latency difference) accounted for significant variance in receptive vocabulary score, R^2 = 0.40, F (3,27) = 5.90, p = .003, $R^2_{Adjusted}$ = 0.33.

The final model estimate of the intercept showed that the mean estimate of receptive vocabulary growth score was 487.19 (*SE* = 1.99) for children with sample average percent of overlapping speech and latency difference, irrespective of group membership. Group membership accounted for significant and unique variance in receptive vocabular score, b = -5.88, *SE* = 2.00, t(27) = -2.94, p = .007, sr^2 = 0.19. In other words, membership in the ASD group was predicted to have an increase in 11.76 points in the child's receptive vocabulary score, compared to membership in the TD group. Holding group membership constant, latency difference accounted for significant and unique variance in receptive vocabulary score at b = 6.17, *SE* = 2.05, t(27) = 3.00, p = .006, sr^2 = 0.20. Meaning, for every one standard deviation increase in latency difference, there is an expected 6.17 point increase in the receptive vocabulary score. Child's percent overlapping speech did not predict receptive vocabulary score (b = 3.35, *SE* = 2.05, t(27) = 1.64, p = .114).

Unlike the results from Research Question 2 primary analyses, both group membership and latency difference significantly and uniquely accounted for variance in receptive vocabulary scores. However, the relationship between the latency difference with receptive vocabulary scores appeared in the opposite direction as predicted: greater difference in latency predicted higher vocabulary scores in this subset of dyads. Lastly, although group membership alone accounted for significant variance in receptive vocabulary, adding the two dyadic coordination variables explained additional significant variances in the outcome, and increased the overall model fit.

Receptive vocabulary as the outcome, tested on dyads with valid receptive vocabulary scores

To test for generalizability of the model, the same regression model with PPVT-5 GSV as the outcome measure and group membership, child's percent overlapping speech, and latency difference as the predictors was tested on a larger sample that included individuals with missing composite languages scores, but with valid receptive vocabulary scores (ASD n = 16, TD n = 24). Note, this sample is different from the sample described for the primary analyses of Research Question 1; one of the cases in the TD group that was removed in the process of group matching for Research Question 1 was included for the purpose of this exploratory analysis.

A bivariate correlation analysis between the outcome and the predictors revealed that difference in latency continued to be positively and significantly related to receptive vocabulary score, r (38) = .31, p = .028. Also, group membership was significantly associated with vocabulary score, r (38) = -0.31, p = .027. Child's percent overlapping speech, r (38) = .08, p = .313, was not correlated with receptive vocabulary.

Again, multiple linear regression with sequential predictor entry was used to test for incremental variance in receptive vocabulary explained by the set of predictors (group, child's percent overlapping speech, and latency difference). For the first model, with group as the only predictor, the model was not significant ($R^2 = 0.094$, $F(1,38) = 3.95$, $p = .054$, $R_{Adjusted} = 0.07$). In the final model, the addition of child's percent overlapping speech and latency difference accounted for additional 15% of variance in receptive vocabulary, $R^2_{change} = 0.15$, $F_{Change}(3,36) = 3.51$, $p = .040$, and the model was significant, $R^2 = 0.24$, $F(3,36) = 3.83$, $p = .018$, $R^2_{Adjusted} = 0.18$. See Table 10.

Table 10.

Multiple Linear Regression with Sequential Predictor Entry for Exploratory Analysis 2; receptive vocabulary as the outcome variable

	Block 1					Block 2				
	R^2_{change}	R^2_{total}	R^2_{adj}	b	sr^2	R^2_{change}	R^2_{total}	R^2_{adj}	b	sr^2
Model Fit	0.09	0.09	0.07			0.15*	0.24*	0.18		
Coefficients										
Intercept				483.71***					483.79***	
Group				-5.04	0.09				-5.45*	0.11
Child's Percent OS									3.81	0.05
Latency Difference									6.49*	0.14

Note. $N = 40$. Block 1 F-change test $df = 1,38$; Block 2 $df = 3, 36$; OS = Overlapping speech; Outcome measure was the PPVT – 5 growth scale value.

* $p < .05$, ** $p < .01$, *** $p < .001$.

In the final model, the estimated intercept showed that dyads with sample average child's percent of overlapping speech and average latency difference, regardless of group membership, predicted 483.79 points on the PPVT-5 growth scale value, $t(36) = 202.99$, $p < .001$. Similar to results from the previous exploratory analysis, group membership significantly accounted for variance in receptive vocabulary score, $b = -5.45$, $SE = 2.39$, $t(36) = -2.28$, $p = .029$, $sr^2 = 0.11$; membership in the ASD group was expected to increase the child's receptive vocabulary score by 10.90 points compared to membership in the TD group. Also, latency difference had a positive relationship with vocabulary score, $b = 6.49$, $SE = 2.53$, $t(36) = 2.57$, $p = .014$, $sr^2 =$

0.14. In other words, for every standard deviation increase in latency difference, there was an estimated increase of 6.43 points in receptive vocabulary score, holding all other variables constant. Child's percent overlapping speech, $b = 3.81$, $SE = 2.53$, $t(36) = 1.51$, $p = .140$, did not predict variance in receptive vocabulary size.

In sum, testing the same regression model on a larger sample size including dyads without composite language scores but with valid receptive vocabulary scores showed latency difference continued to be a unique and significant factor for children in the ASD and TD groups combined. Furthermore, the unexpected positive relationship between latency difference and language abilities persisted in all three regression analyses. In both exploratory analyses, group membership was a significant and unique predictor of receptive vocabulary size. However, adding the dyadic coordination variables accounted for larger variance in the receptive vocabulary score, leading to a better model fit and significance.

DISCUSSION

The dyadic coordination of behaviors, such as vocalizations and pauses between caregivers and their children are evident since the early stages of development (Leclère et al., 2014; Northrup & Iverson, 2020). This dyadic interaction is thought to give rise to outcomes across psychosocial domains, including the development of language (Harrist & Waugh, 2002; Jaffe et al., 2001). Moreover, differences in dyadic interaction in children with neurodevelopmental disorders may eventually serve as potential diagnostic markers or treatment targets for speech and language therapy.

Considering evidence in the literature that dyadic coordination of vocalization and pauses in infant caregiver dyads are associated with language abilities in young children with and without neurodevelopmental disorders (Condon & Sander, 1974; Kellerman et al., 2020; Northrup & Iverson, 2015; Zampella et al., 2020), it was speculated that dyadic coordination variables such as child's percentage of overlapping speech and quantitative estimates of the caregiver and child's alignment with each other in timing of vocal interaction, may be associated with language skills in older autistic children. Specifically, results from an infant-caregiver dyad study also indicated that dyadic coordination variables such as infant's percent of overlapping speech and the difference in latency between the infant and the caregiver were related to language delays and later language abilities, respectively. (Northrup & Iverson, 2015). Indeed, in our preliminary study, results indicated that child's percent of overlapping speech in a structured interactive task with an examiner was significantly and negatively associated with vocabulary size in school-age autistic children (Choi & Kover, under review).

The current study extended the literature on dyadic coordination by examining children's interactions with their caregivers during the school-age years, in a naturalistic environment, and in relation to specific aspects of language (i.e., overall language ability and receptive vocabulary). Child and caregiver's percent overlapping speech and latency difference (i.e., the difference between the child's and caregiver's mean length of between-speaker pauses), were similar to the variables previously reported in the literature (Gratier, 2003; Gratier et al., 2015; Northrup & Iverson, 2015). The difference in percent overlapping speech between the caregiver and the child was newly explored in this current study. The combination of using documented and novel dyadic coordination variables provided context to the current findings, as well as expanded on methods in capturing dyadic coordination variables.

Further, the relationship between the child and caregiver's conversational state variables, such as frequency and mean duration of vocalization and pauses, were analyzed to fill the knowledge gap in dyadic interaction between older autistic children and their caregivers engaged in parent-child-play sessions in a naturalistic environment. Findings from the current study regarding similarities and differences in the conversational state variables between dyads with autistic children compared to dyads with typically developing children are parallel to findings in the literature on younger autistic children and expands on findings from adolescents with autism (Feldstein et al., 1982; Northrup & Iverson, 2015). Moreover, these quantitative estimates for dyads and the alignment between child and caregiver in and of itself has potential for future studies, especially around culture, linguistic exposure, and caregiver child characteristics.

DYADIC COORDINATION BETWEEN CHILD AND CAREGIVER

The initial set of analyses provided context and understanding of the dyadic coordination constructs in the current study (given older children, a remotely collected parent-child play task, etc.), as well as situated the current observations of dyadic coordination in the literature. In the current study, conversational state variables that pertained to dyadic coordination variables (i.e., frequency and duration of latency and frequency of overlapping speech) were significantly and positively correlated between children and their caregivers in both the ASD and TD dyads. In other words, high frequency and duration of overlapping speech by caregivers during parent-child play, was associated with high frequency and duration of overlapping speech observed in their child. The duration of overlapping speech was also significantly and positively correlated in TD dyads, however, did not reach significance in the ASD dyads. Although the duration of overlapping speech in ASD dyads were not significantly correlated, these correlations suggest overall that there is some alignment in the tendency to produce overlapping vocalizations between children and their caregivers, regardless of ASD diagnostic status.

Positively correlated frequency and duration of latency (i.e., between-speaker pauses) might reflect shared vocal interaction styles for each dyad; a slower pace with more gaps for some dyads and shorter gaps for other dyads. These results are in alignment with the metatheoretical concept "reciprocity" where each partner contributes equally in quantity and quality to the interaction (Provenzi et al., 2018). From the current study, it appears that dyads from both the ASD and TD group engaged in reciprocal coordination of the frequency and duration of overlapping speech and between-speaker pauses.

POTENTIAL GROUP DIFFERENCES

The first primary research question asked whether dyadic coordination differed between the ASD and TD groups. There were no detectable group differences in the dyadic coordination variables (child's and caregiver's percent overlapping speech, latency difference) or the exploratory dyadic coordination variable (difference in percent overlapping speech). However, a group difference was detected in two of the conversational state variables: child's frequency of vocalizations and frequency of intrapersonal pauses (i.e., within-speaker pauses) were lower in the ASD group compared to the TD group. These results are in alignment with findings from Warlaumont et al. (2014), where researchers observed reduced frequency of vocalization and weakened social feedback loops in dyads of 8- to 48-month-old autistic children with adults in the home environment. Although the current study was not focused on analyzing social feedback loops between dyads, evidence of positively correlated conversational state variables such as frequency and duration of overlapping speech and between-speaker pauses point towards similar reciprocal interactions between older autistic children and their caregivers in a natural environment. As such, it is possible that despite reduced frequency of the autistic child's vocalization, the reciprocal interaction between child and caregiver's vocal interactions continues into later stages of development.

In a study that examined pauses in autistic adolescents engaged in conversation with the caregivers or the examiners, the adolescents' familiarity to the conversation partner played a role in the duration of between- and within-speaker pauses (Feldstein et al., 1982). However, the study did not examine group differences between autistic adolescents and typically developing adolescents in conversation with familiar or unfamiliar adults. The current study shows that although there is reduced frequency of intrapersonal gaps in autistic children compared to typically developing children, there were no differences detected in the *duration* of between- or within-speaker pauses, or latency difference, when they interacted with their caregivers in the home environment. At the same time, there was a large but nonsignificant effect with increased intrapersonal gap frequency in caregivers of autistic participants. This could relate to allowing time for a child response, but then continuing with a next vocalization if the child did not respond, although this interpretation would need to be verified in the data. Taken together, it is possible that autistic children engaged with familiar adults in a familiar environment show minimal group differences from TD dyads in the frequency and duration of pauses, again suggesting some potential commonalities in caregiver-child vocalization behaviors between dyads with autistic and neurotypical children.

Beyond group differences found in the frequency of vocalization and intrapersonal gaps, no other differences were observed in conversational state variables at large (e.g., between-speaker pauses, overlapping speech). The lack of differences observed in not only the child's dyadic coordination and conversational state variables, but also that of the caregivers, indicate that dyads appear to engage similarly, at least in terms of some measures, in parent-child-play sessions regardless of the child's diagnostic status.

RELATIONSHIP BETWEEN DYADIC COORDINATION AND CHILD LANGUAGE ABILITY

The second primary research question tested the relationship between child's percent of overlapping speech and latency difference with language abilities. First, the set of predictors (group, child's percent overlapping speech, and latency difference) did not account for significant variance in concurrent composite language score (i.e., CELF-P3 growth scale values). Initially, the bivariate correlation analysis conducted before the regression analysis showed that latency difference was significantly and positively associated with the child's composite language score. However, this relationship did not hold in the regression model, as the overall model did not reach significance. These results were unexpected, as a previous study examining infant-caregiver dyads reported that latency difference predicted later composite language scores (Northrup & Iverson, 2015).

One potential explanation of the current finding is that an existing relationship between one or more of the predictors with the outcome was not detected in the current analysis due to limited sample size. This regression model was tested on a group of dyads with valid composite language scores, limiting the total sample size to n = 31. On the other hand, it is also possible that the relationship between dyadic coordination variables and concurrent language abilities do not present the same way as observed in the younger population compared to older children with more advanced language abilities. It may be that dyadic coordination could have a cascading or compounding effect on speech and language abilities, such that later in development other aspects of spoken language or social interaction become more salient and impactful on language development. Future studies should consider investigating potential compounding or downstream factors related to dyadic coordination variables on later language abilities.

To further test the relationship between the dyadic coordination variables with language abilities, the same set of predictors was regressed on receptive vocabulary scores (i.e., PPVT-5 growth scale values) in two exploratory analyses. The first exploratory regression model was tested on the same sample of dyads with complete composite language scores and receptive vocabulary scores (ASD n = 11, TD n = 20). The results showed that latency difference predicted receptive vocabulary score, keeping group membership constant. To test for generalizability of the model, the sample size was expanded to include dyads without composite language scores, but with valid receptive vocabulary scores (ASD n = 14; TD n = 26). Results indicated that latency difference continued to significantly and uniquely predict variance in receptive vocabulary scores, regardless of group membership. Notably, the relationship between latency difference and receptive vocabulary score from both analyses was positive, which was the opposite direction of effect expected based on a preliminary study and the existing literature.

One potential explanation for this unexpected result is that the observed child's percent of overlapping speech and latency difference with language abilities may vary according to the context and environment. Based on the findings from Jaffe et al. (2001), the dyadic coordination variables may be impacted by the child's familiarity with the speaker or the environment. Further, coordination patterns observed in infant-unfamiliar adult dyads were more predictive of cognitive skills, whereas coordination patterns observed in infant-familiar adult dyads were better predictors of attachment levels. Similarly, the child's percent overlapping speech observed in a structured task with an unfamiliar adult may be significantly and negatively associated with concurrent vocabulary skills (under review; Choi & Kover, 2021); however, the child's percent overlapping speech observed in a semi-structured play session with a familiar adult in the home environment may not be a reliable predictor of concurrent vocabulary skills in older children. Also, the relationship between latency difference and language abilities may present differently in infant-caregiver dyads (Northrup & Iverson, 2015) compared to school-age children-caregiver dyads; while increased difference in latency in infant-caregiver dyads indicate asynchronous interaction predicting lower composite language scores, the same observation may not indicate a difficulties with language development for a school-age child engaged in play with their caregivers.

The positive relationship between latency difference and language abilities observed in school-aged children may be indicative of different types of dyadic coordination utilized by dyads across various contexts (e.g., structured tasks, play sessions) and across developmental time. It is possible that latency (i.e., space between vocalizations) observed between school-age autistic children and examiner engaged in a structured task is reciprocal in nature, where each conversation partner is expected to contribute equally in quantity and quality of the interaction (Northrup & Iverson, 2015; Provenzi et al., 2018), including the duration of between-speaker pauses. Conversely, latency observed between school-age autistic children and their caregivers engaged in parent-child-play session may be mutual, where there are clear differences in the quality or quantity of their contributions to the interaction (Provenzi et al., 2018; Roe & Drivas, 1997; Savelkoul et al., 2007), but the difference could represent an environment that is not detrimental to language learning. Perhaps an interaction that is reciprocal (that is, equal in quality and quantity) in the context of natural parent-child-play session is indicative of a more rigid interaction that is not conducive to language learning, thereby associated with lower language abilities, in school-age children. The current findings provide an impetus for further investigating the meta-theoretical concepts of "mutuality" and "reciprocity" across developmental time and context in autistic children (Provenzi et al., 2018).

SUMMARY OF FINDINGS

In sum, the reciprocal nature of certain dyadic coordination variables (i.e., frequency and mean duration of latency and overlapping speech) between school-age children and their caregivers were observed regardless of diagnostic status. While group differences were observed in the frequency of vocalization and intrapersonal gaps, the dyadic coordination variables at large were not different between the ASD and TD dyads. Lastly, latency difference (in other words, the difference in the mean length of pauses ended by the child or caregiver) was a significant and unique predictor of receptive vocabulary size in school-age children. The positive relationship between difference in latency and receptive vocabulary suggests that the nature of latency difference observed in parent-child-play sessions may be different from those measured in structured language learning tasks with unfamiliar adult and in older school-age children with a familiar interaction partner.

LIMITATIONS AND FUTURE DIRECTIONS

Despite rigorous protocols for implementing an online study, there remained some technical difficulties and natural disturbances that may have impacted the quality of the data. For instance, there was one parent-child-play session that could not be used due to internet connection issues and poor audio quality. Some parent-child-play sessions occurred in rooms with poor acoustics, which may have impacted the coders abilities in segmenting the beginning and ending of utterances. In addition, several parent-child-play sessions included interruptions from the home environment that impacted the interaction between the primary caregiver and the child (e.g., pets or other household members intruding). Nevertheless, there is evidence that language samples collected from online video-calling software (e.g., Skype, Google Hangouts, Face Time) and devices (e.g., smartphone, tablet, laptop, desktop) do not differ in quality compared to language samples collected in person (Manning, Harpole, Harriott, Postolowicz, & Norton, 2020). Furthermore, 20% of the parent-child-play sessions were double coded by two independent and trained coders to ensure reliability and the quality of data. Altogether, every effort was made to ensure the data extracted from the online study would be credible, despite these limitations. However, future studies that examine language in the natural environment should consider the use of higher quality voice recording devices that can be easily accessed by the caregivers (e.g., mailed to families' homes prior to the study session).

Standard procedures were followed for the administration of online language assessments (PPVT-5, CELF- P3); however, there were limitations to readily accessing assessments that provide further information about the participant characteristics (e.g., autism symptom severity and nonverbal cognition). While the original in-person study design of the current study included direct assessment of the child's autism symptom severity using the Autism Diagnostic Observation Schedule (Lord et al., 2000) and nonverbal cognition using the Leiter-3 (Roid, Miller, Pomplun, & Koch, 2013), they could not be administered when the research pivoted to a remote online study design due to the COVID-19 pandemic.

A parent questionnaire (i.e., SRS – 2) was used instead of a direct assessment to quantify the autism symptom severity of the current sample; nevertheless, 5 out of 20 autistic children had missing SRS-2 scores. As such, there is a limitation in that, the current findings generalize to autistic children with community diagnoses, rather than research classifications of autism. Alternative tools that are suitable for remote administration are becoming much more prevalent and well-studied, including the Brief Observation of Symptoms of Autism (BOSA; Dow et al., 2022).

The current study matched the groups using PPVT-5 scores, as there was theoretical motivation in finding group differences in dyadic coordination variables among children with similar receptive vocabulary scores. However, additionally being able to match the groups on nonverbal cognition would have been an asset. Perhaps even more importantly, including nonverbal cognition in models predicting language would have been ideal given the known association between nonverbal developmental level and language in autism (Ellis Weismer & Kover, 2015). In addition, a chronological age matched comparison group would also have been insightful.

Lastly, the participant sample lacked diversity in many senses, including in multicultural and multilingual backgrounds and diverse socio-economic statuses. As such, there is a limitation to the generalizability of our results. Future research should focus on dyadic coordination in multilingual dyads and multicultural dyads to better understand how linguistic and cultural variation impact social interaction, vocal behaviors, and language development in autism. Studying these factors over developmental time in a longitudinal study would be especially useful. Further, there is evidence that the gender of the caregiver may influence dyadic coordination variables (Feldman, 2007; Leclère et al., 2014; Savelkoul et al., 2007; Trevarthen & Daniel, 2005). The current study did not balance or control the gender of the caregivers and assumed the caregivers' gender identity based on visual inspection of the parent-child-play sessions. As such, future studies investigating dyadic coordination between child-caregiver dyads should incorporate a more diverse population of dyads and account for linguistic and cultural diversity in their data analyses.

In conclusion, dyadic coordination of vocalization and pauses appear to play a role in child's language abilities, in particular, concurrent vocabulary size. However, there remains many unanswered questions about dyadic coordination variables. Future studies should investigate how specific dyadic coordination variables, such as latency differences or child's percent of overlapping speech observed in parent-child-play sessions may differ given the context and structure of an interaction. Further, understanding the relationship and its unique contributions to language abilities may potentially lead to language treatment strategies or targets.

References

Agt, H., Verhoeven, L., van den Brink, G., & de Koning, H. (2011). The impact on socio-emotional development and quality of life of language impairment in 8-year-old children. *Developmental Medicine and Child Neurology*, *53*(1), 81–88. https://doi.org/10.1111/j.1469-8749.2010.03794.x

Bayley, N. (1993). Bayley scales of infant development. (P. Corporation, Ed.). New York: New York : Psychological Corp.

Beebe, B., Alson, D., Jaffe, J., Feldstein, S., & Crown, C. (1988). Vocal Congruence in Mother-Infant Play. *Journal of Psycholinguistic Research*, *17*(3), 245–259. https://doi.org/10.1007/BF01686358

Boersma, P., & Weenink, D. (2019). Praat: doing phonetics by computer [Computer program]. *Glot International*.

Chan, C. Y. Z., Williams, K., May, T., Wan, W. H., & Brignell, A. (2023). Is language ability associated with behaviors of concern in autism? A systematic review. *Autism Research*, *16*(2), 250–270. https://doi.org/10.1002/aur.2855

Choi, E. A., & Kover, S. T. (n.d.). Social and Linguistic Correlates of Vocabulary Size in Autism: Overlapping Speech and Phonological Memory. *Journal of Autism and Developmental Disorders*.

Coales, C., Heaney, N., Ricketts, J., Dockrell, J. E., Lindsay, G., Palikara, O., & Charman, T. (2019). Health-related quality of life in children with autism spectrum disorders and children with developmental language disorders. *Autism & Developmental Language Impairments*, *4*, 239694151985122. https://doi.org/10.1177/2396941519851225

Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, *20*(1), 37–46. https://doi.org/10.1177/001316446002000104

Condon, W. S., & Sander, L. W. (1974). Neonate Movement Is Synchronized with Adult Speech: Interactional Participation and Language Acquisition. *Science (American Association for the Advancement of Science)*, *183*(4120), 99–101. https://doi.org/10.1126/science.183.4120.99

Constantino, J., & Gruber, C. (2012). Social Responsiveness Scale (SRS): Manual. *Los Angeles, CA: Western Psychological Services*.

Donnelly, S., & Kidd, E. (2021). The Longitudinal Relationship Between Conversational Turn-Taking and Vocabulary Growth in Early Language Development. *Child Development*, *92*(2), 609–625. https://doi.org/https://doi.org/10.1111/cdev.13511

Dunn, D. M. (2019). PPVT-5 : Peabody Picture Vocabulary Test. Form A. (publisher Pearson (Firm), Ed.), *Peabody Picture Vocabulary Test. Form A*. Bloomington, MN.: Bloomington, MN. : Pearson.

Dunn, L. M., & Dunn, D. M. (2007). *PPVT-4 : Peabody picture vocabulary test. Peabody picture vocabulary test*.

Ellis Weismer, S., & Kover, S. T. (2015). Preschool language variation, growth, and predictors in children on the autism spectrum. *Journal of Child Psychology and Psychiatry and Allied Disciplines*. https://doi.org/10.1111/jcpp.12406

Farran, L. K., Yoo, H., Lee, C.-C., Bowman, D. D., & Oller, D. K. (2019). Temporal Coordination in Mother–Infant Vocal Interaction: A Cross-Cultural Comparison. *Frontiers in Psychology*, *10*.

Feeney, R., Desha, L., Khan, A., & Ziviani, J. (2017). Contribution of speech and language difficulties to health-related quality-of-life in Australian children: A longitudinal analysis. *International Journal of Speech Language Pathology*, *19*(2), 139–152. https://doi.org/10.3109/17549507.2016.1151935

Feldman, R. (2007). Parent-infant synchrony and the construction of shared timing; physiological precursors, developmental outcomes, and risk conditions. *Journal of Child Psychology and Psychiatry and Allied Disciplines*. https://doi.org/10.1111/j.1469-7610.2006.01701.x

Feldman, R. (2010). The relational basis of adolescent adjustment: trajectories of mother-child interactive behaviors from infancy to adolescence shape adolescents' adaptation. *Attachment & Human Development*, *12*(1–2), 173–192. https://doi.org/10.1080/14616730903282472

Feldman, R., Greenbaum, C. W., & Yirmiya, N. (1999). Mother-Infant Affect Synchrony as an Antecedent of the Emergence of Self-Control. *Developmental Psychology*, *35*(1), 223–231. https://doi.org/10.1037/0012-1649.35.1.223

Feldstein, S., Konstantareas, M., Oxman, J., & Webster, C. D. (1982). The chronography of interactions with autistic speakers: An initial report. *Journal of Communication Disorders*, *15*(6), 451–460. https://doi.org/10.1016/0021-9924(82)90018-1

Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., Reznick, J. S., & Bates, E. (2007). MacArthur-Bates Communicative Development Inventories (2nd ed.). Paul H. Brookes Publishing Co.

Frazier, T. W., Whitehouse, A. J. O., Leekam, S. R., Carrington, S. J., Alvares, G. A., Evans, D. W., … Uljarević, M. (2023). Reliability of the Commonly Used and Newly-Developed Autism Measures. *Journal of Autism and Developmental Disorders*. https://doi.org/10.1007/s10803-023-05967-y

Gratier, M. (2003). Expressive timing and interactional synchrony between mothers and infants: Cultural similarities, cultural differences, and the immigration experience. *Cognitive Development*. https://doi.org/10.1016/j.cogdev.2003.09.009

Gratier, M., Devouche, E., Guellai, B., Infanti, R., Yilmaz, E., & Parlato-Oliveira, E. (2015). Early development of turn-taking in vocal interaction between mothers and infants. *Frontiers in Psychology*, *6*, 1167.

Gray, S. (2004). Word Learning by Preschoolers With Specific Language Impairment: Predictors and Poor Learners. *Journal of Speech, Language, and Hearing Research @BULLET Kiernan & Gray Rice*, *47*, 1117–1132. https://doi.org/10.1044/1092-4388(2004/083)

Gray, S. (2006). The relationship between phonological memory, receptive vocabulary, and fast mapping in young children with specific language impairment. *Journal of Speech, Language, and Hearing Research : JSLHR*, *49*(5), 955–969. https://doi.org/10.1044/1092-4388(2006/069)

Harrist, A. W., & Waugh, R. M. (2002). Dyadic synchrony: Its structure and function in children's development. *Developmental Review*. https://doi.org/10.1016/S0273-2297(02)00500-2

Healey, D. M., Gopin, C. B., Grossman, B. R., Campbell, S. B., & Halperin, J. M. (2010). Mother-child dyadic synchrony is associated with better functioning in hyperactive/inattentive preschool children. *Journal of Child Psychology and Psychiatry and Allied Disciplines*. https://doi.org/10.1111/j.1469-7610.2010.02220.x

Hill, A. P., van Santen, J., Gorman, K., Langhorst, B. H., & Fombonne, E. (2015). Memory in language-impaired children with and without autism. *Journal of Neurodevelopmental Disorders*, *7*(1), 19. https://doi.org/10.1186/s11689-015-9111-z

Hoch, J. E., Ossmy, O., Cole, W. G., Hasan, S., & Adolph, K. E. (2021). "Dancing" Together: Infant–Mother Locomotor Synchrony. *Child Development*, *92*(4), 1337–1353. https://doi.org/10.1111/cdev.13513

Hove, M. J., & Risen, J. L. (2009). IT'S ALL IN THE TIMING: INTERPERSONAL SYNCHRONY INCREASES AFFILIATION. *Social Cognition*, *27*(6), 949–961. https://doi.org/10.1521/soco.2009.27.6.949

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*.

Jasnow, M., Crown, C. L., Feldstein, S., Taylor, L., Beebe, B., & Jaffe, J. (1988). Coordinated Interpersonal Timing of Down-Syndrome and Nondelayed Infants with Their Mothers: Evidence for a Buffered Mechanism of Social Interaction. *The Biological Bulletin (Lancaster)*, *175*(3), 355–360. https://doi.org/10.2307/1541726

Kellerman, A. M., Schwichtenberg, A. J., Abu-Zhaya, R., Miller, M., Young, G. S., & Ozonoff, S. (2020). Dyadic Synchrony and Responsiveness in the First Year: Associations with Autism Risk. *Autism Research*. https://doi.org/10.1002/aur.2373

Kwok, E., Feiner, H., Grauzer, J., Kaat, A., & Roberts, M. Y. (2022). Measuring Change During Intervention Using Norm-Referenced, Standardized Measures: A Comparison of Raw Scores, Standard Scores, Age Equivalents, and Growth Scale Values From the Preschool Language Scales-Fifth Edition. *Journal of Speech, Language, and Hearing Research*, *65*(11), 4268–4279. https://doi.org/10.1044/2022_JSLHR-22-00122

Leclère, C., Viaux, S., Avril, M., Achard, C., Chetouani, M., Missonnier, S., & Cohen, D. (2014). Why synchrony matters during mother-child interactions: A systematic review. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0113571

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Leventhal, B. L., Dilavore, P. C., … Rutter, M.

(2000). The Autism Diagnostic Observation Schedule-Generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders*. https://doi.org/10.1023/A:1005592401947

Lowinger, S. (1999). Infant irritability and early mother-infant reciprocity patterns. *Infant and Child Development*, *8*(2), 71–84. https://doi.org/10.1002/(SICI)1522-7219(199906)8:2<71::AID-ICD191>3.0.CO;2-1

Manning, B. L., Harpole, A., Harriott, E. M., Postolowicz, K., & Norton, E. S. (2020). Taking language samples home: Feasibility, reliability, and validity of child language samples conducted remotely with video chat versus in-person. *Journal of Speech, Language, and Hearing Research*. https://doi.org/10.1044/2020_JSLHR-20-00202

McKnight, P. E., & Najab, J. B. T.-T. C. E. of P. (2010). Mann-Whitney U Test. In I. B. e & W. E. Craighead (Eds.) (4th ed. IB, Vol. 3, pp. 960–961). Hoboken, NJ: Wiley. Retrieved from https://link.gale.com/apps/doc/CX1942000537/GVRL?u=wash_main&sid=bookmark-GVRL&xid=f27c95f0

Mullen, E. M. (1995). Mullen Scales of Early Learning, AGS Edition: Manual and Item Administrative Books. *American Guidance Services, Inc*.

Murray, L., De Pascalis, L., Bozicevic, L., Hawkins, L., Sclafani, V., & Ferrari, P. F. (2016). The functional architecture of mother-infant communication, and the development of infant social expressiveness in the first two months. *Scientific Reports*, *6*(1), 39019. https://doi.org/10.1038/srep39019

Nadig, A., & Mulligan, A. (2017). Intact non-word repetition and similar error patterns in language-matched children with autism spectrum disorders: A pilot study. *Journal of Communication Disorders*, *66*, 13–21. https://doi.org/10.1016/j.jcomdis.2017.03.003

Nomikou, I., Koke, M., & Rohlfing, K. J. (2017). Verbs in Mothers' Input to Six-Month-Olds: Synchrony between Presentation, Meaning, and Actions Is Related to Later Verb Acquisition. *Brain Sciences*, *7*(5). https://doi.org/10.3390/brainsci7050052

Norbury, C. F., Griffiths, H., & Nation, K. (2010). Sound before meaning: Word learning in autistic disorders. *Neuropsychologia*, *48*(14), 4012–4019. https://doi.org/10.1016/j.neuropsychologia.2010.10.015

Northrup, J. B. J. B., & Iverson, J. M. J. M. (2020). Multi-Modal Coordination of Vocal and Gaze Behavior in Mother-Infant Dyads Across the First Year of Life. *Infancy*, *25*(6), 952–972. https://doi.org/10.1111/infa.12369

Northrup, J., & Iverson, J. (2015). Vocal Coordination During Early Parent-Infant Interactions Predicts Language Outcome in Infant Siblings of Children with Autism Spectrum Disorder. *Infancy*. https://doi.org/10.1111/infa.12090

Northrup, J., & Iverson, J. (2020). The development of mother–infant coordination across the first year of life. *Developmental Psychology*, *56*(2), 221–236. https://doi.org/10.1037/dev0000867

Pearson Inc. (2023). Q-global. Retrieved from https://www.pearsonassessments.com/store/usassessments/en/Store/Professional-Assessments/Q-global-Web-based-Administration%2C-Scoring%2C-and-Reporting/p/100000680.html?tab=product-details

Provenzi, L., di Minico, G. S., Giusti, L., Guida, E., & Müller, M. (2018). Disentangling the dyadic dance: Theoretical, methodological and outcomes systematic review of mother-infant dyadic processes. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2018.00348

Rauchbauer, B., & Grosbras, M. H. (2020). Developmental trajectory of interpersonal motor alignment: Positive social effects and link to social cognition. *Neuroscience & Biobehavioral Reviews*, *118*, 411–425. https://doi.org/10.1016/J.NEUBIOREV.2020.07.032

Roe, K. V, & Drivas, A. (1997). RECIPROCITY IN MOTHER-INFANT VOCAL INTERACTIONS: Relationship to the Quantity of Mothers' Vocal Stimulation. *American Journal of Orthopsychiatry*, *67*(4), 645–649. https://doi.org/10.1037/h0080262

Roid, G., Miller, L., Pomplun, M., & Koch, C. (2013). Leiter International Performance Scale-Third Edition (Leiter-3).

Ryant, N., Yuan, J., & Liberman, M. (2013). Automating phonetic measurement: The case of voice onset time. *The Journal of the Acoustical Society of America*, *133*(5), 3608. https://doi.org/10.1121/1.4806713

Savelkoul, E. M., Zebrowski, P. M., Feldstein, S., & Cole-Harding, S. (2007). Coordinated interpersonal timing in the conversations of children who stutter and their mothers and fathers. *Journal of Fluency Disorders*, *32*(1), 1–32. https://doi.org/10.1016/j.jfludis.2006.12.001

Seidl, A., Cristia, A., Soderstrom, M., Ko, E. S., Abel, E. A., Kellerman, A., & Schwichtenberg, A. J. (2018). Infant–mother acoustic–prosodic alignment and developmental risk. *Journal of Speech, Language, and Hearing Research*. https://doi.org/10.1044/2018_JSLHR-S-17-0287

Siegel, L. S. (1989). A reconceptualization of prediction from infant test scores. In M. H. Bornstein & N. A. Krasnegor (Eds.), *Stability and continuity in mental development: Behavioral and biological perspectives* (pp. 89–103). Hillsdale, N.J.: Erlbaum Associates.

Skuban, E. M., Shaw, D. S., Gardner, F., Supplee, L. H., & Nichols, S. R. (2006). The correlates of dyadic synchrony in high-risk, low-income toddler boys. *Infant Behavior & Development*, *29*(3), 423–434. https://doi.org/10.1016/j.infbeh.2006.02.004

Smith, J., Chetcuti, L., Kennedy, L., Varcin, K. J., Slonims, V., Bent, C. A., … Hudry, K. (2023). Caregiver sensitivity predicts infant language use, and infant language complexity predicts caregiver language complexity, in the context of possible emerging autism. *Autism Research*, *16*(4), 745–756. https://doi.org/10.1002/aur.2879

Tabachnick, B. G., & Fidell, L. S. (2012). *Using multivariate statistics (6th ed.). New York: Harper and Row*. https://doi.org/10.1037/022267

Trevarthen, C., & Daniel, S. (2005). Disorganized rhythm and synchrony: Early signs of autism and Rett syndrome. In *Brain and Development* (Vol. 27, pp. S25–S34). Elsevier.

https://doi.org/10.1016/j.braindev.2005.03.016

VideoLAN. (n.d.). VLC media player. VideoLAN. Retrieved from https://www.videolan.org/vlc/

Wan, M. W., Green, J., Elsabbagh, M., Johnson, M., Charman, T., & Plummer, F. (2013). Quality of interaction between at-risk infants and caregiver at 12-15 months is associated with 3-year autism outcome. *Journal of Child Psychology and Psychiatry*, *54*(7), 763–771. https://doi.org/10.1111/jcpp.12032

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*. https://doi.org/10.1177/0956797614531023

Weismer, S. E., Tomblin, J. B., Zhang, X., Buckwalter, P., Chynoweth, J. G., & Jones, M. (2000). Nonword Repetition Performance in School-Age Children with and Without Language Impairment. *Journal of Speech, Language, and Hearing Research*. https://doi.org/10.1044/jslhr.4304.865

Welkowitz, J., Cariffe, G., & Feldstein, S. (1976). Conversational Congruence as a Criterion of Socialization in Children. *Child Development*, *47*(1), 269–272. https://doi.org/10.1111/j.1467-8624.1976.tb03421.x

Welkowitz, J., & Kuc, M. (1973). Interrelationships among warmth, genuineness, empathy, and temporal speech patterns in interpersonal interaction. *Journal of Consulting and Clinical Psychology*, *41*(3), 472–473. https://doi.org/10.1037/h0035328

Wiig, E. H. (2020). CELF preschool 3 : clinical evaluation of language fundamentals - preschool. *CELF Preschool 3 : Clinical Evaluation of Language Fundamentals - Preschool*. Bloomington, Minnesota: Pearson, Inc.

Williams, K. T. (2007). Expressive Vocabulary Test (2nd Edition). *J. Am. Acad. Child Adolesc. Psychiatry*.

Yirmiya, N., Gamliel, I., Pilowsky, T., Feldman, R., Baron-Cohen, S., & Sigman, M. (2006). The development of siblings of children with autism at 4 and 14 months: Social engagement, communication, and cognition. *Journal of Child Psychology and Psychiatry and Allied Disciplines*. https://doi.org/10.1111/j.1469-7610.2005.01528.x

Zampella, C., Csumitta, K., Simon, E., & Bennetto, L. (2020). Interactional Synchrony and Its Association with Social and Communication Ability in Children With and Without Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*.