Identifying Vocal-Motor Behaviors of Joint Engagement in Children with Autism Spectrum Disorder and Their Parents

Natasha M. Yamane

Teachers College, Columbia University

Author Note

Natasha M. Yamane, Department of Counseling and Clinical Psychology, Teachers College, Columbia University.

I thank Dr. Sylvie Goldman in the Department of Neurology at the Columbia University Irving Medical Center for her dedication and mentorship, Aston K. McCullough in the Department of Biobehavioral Sciences at Teachers College, Columbia University for his feedback on all matters of statistics and computer programming, and Dr. Letitia Naigles in the Department of Psychology at the University of Connecticut for her support and generosity in sharing her data for this project.

Correspondence concerning this paper should be addressed to Natasha M. Yamane, Department of Counseling and Clinical Psychology, Teachers College, Columbia University, New York, NY, 10027. Email: nmy2103@tc.columbia.edu

Table of Contents

Abstract

Impairments in joint engagement (JE), the triadic arrangement between a parent and a child around a shared object or event, have been vastly studied as a hallmark of children with autism spectrum disorder (ASD). However, a majority of existing work on JE has used primarily global measures of social behaviors derived from prior work on joint attention skills. To address this limitation, we developed the Vocal-Motor Coding System (VMCS), a novel coding system that integrates motoric variables of Proximity, Orientation, and Object–Touch, as well as vocal variables of Loudness and Rhythmicity (i.e., vocalization, pause, and switching pause). The VMCS was applied to code a joint-attention task completed by 20 parent-child dyads, including 10 with typical development (TD) and 10 with ASD. The criterion validity of the VMCS was assessed against an established coding system for JE, revealing a number of significant correlations ($p < 0.05$) between motoric behaviors and child engagement states and between vocal behaviors and parental attention-directing strategies. Although no significant differences were found between vocal-motor behaviors of dyads with TD and ASD, we found strong, positive associations within the ASD group among the frequencies and durations of dyadic motoric behaviors with parental vocal behaviors. Specifically, close proximity was strongly related to parental loudness ($\rho = -0.90$) and switching pauses of parents ($\rho = 0.89$) and children ($\rho = 0.93$). Findings support the viability of the VMCS as an instrument for coding JE using discrete vocal-motor measures, and point toward its utility in characterizing strategies used by parents to achieve JE with their children.

Keywords: autism spectrum disorder, joint engagement, dyadic behaviors, parent-child interactions, motor behaviors, vocal rhythm

Identifying Vocal-Motor Behaviors of Joint Engagement in Parents and Their Children with Autism Spectrum Disorder

Joint engagement, the triadic arrangement between a parent and a child around a shared object or event, provides an important context for the development of social cognition (Tomasello, 1995). Impairments in joint engagement are commonly identified as a primary marker of children with autism spectrum disorder (Mundy & Newell, 2007; Mundy, Sullivan, & Mastergeorge, 2009), a neurodevelopmental disorder of early onset characterized by social deficits. Since early work on social interactions between parents and children with autism spectrum disorder (Bakeman & Adamson, 1984; Sigman, Mundy, Sherman, & Ungerer, 1986), a large body of research has identified a number of early behavioral indicators that are essential to fostering dyadic engagement between parents and their children. Specifically, these indicators comprise verbal behaviors, including response to name and pre-verbal vocalizations such as cooing and babbling, as well as nonverbal behaviors, such as eye gaze and communicative gestures that altogether constitute an individual's attention on his or her social partner and a mutual object or event. In infancy, the practice of these behaviors through the cortical activation of a distributed attention network (Mundy & Newell, 2007) contributes to the infant's development of socio-cognitive awareness. A number of studies have shown that young children with autism spectrum disorder (ASD) demonstrate these behaviors in their capacity to respond to joint attention with a social partner, but compared to children with typical development (TD), are less adept at using these behaviors in the context of initiating joint attention (Charman, 2003; Dawson et al., 2004; Mundy, Sigman, Ungerer, & Sherman, 1986a). Consequently, these behavioral indicators have come to serve as discriminators of ASD in young children in various approaches to early detection (Sigman, Dijamco, Gratier, & Rozga, 2004), including the "gold standard" diagnostic instrument, the *Autism Diagnostic Observation Schedule* (ADOS; Lord et al., 2000). They have also served as measures of social reciprocity in studies of parent-child interactions (De Barbaro, Johnson, Forster, & Deák, 2016; Patterson, Elder, Gulsrud, & Kasari, 2014) that have integrated our current understanding of joint attention in the context of assessing language development (Adamson, Bakeman, & Suma, 2017; Bottema-Beutel, Yoder, Hochman, & Watson, 2014), attachment (Naber et al., 2007), emotional regulation (Gulsrud, Jahromi, & Kasari, 2010; Ting & Weiss, 2017; Zhou & Yi, 2014), object manipulation (Zukow-Goldring & Arbib, 2007), as well as parent-mediated interventions (Chiang, Chu, & Tsung-Chin Lee, 2016; Girolametto, Verbey, & Tannock, 1994; Siller, Swanson, Gerber, Hutman, & Sigman, 2014).

Joint Attention or Joint Engagement?

Research to date on behaviors of social reciprocity, or "jointness" (Hobson & Hobson, 2011) in children with autism spectrum disorder has appeared to use "joint attention" interchangeably with "joint engagement." Before the current review of the literature, a distinction between *joint attention* (JA) and *joint engagement* (JE) is warranted. Hobson and Hobson (2011) have differentiated the two concepts by the contexts in which they are observed and measured. To study attention, they argue, we must consider the type of engagement in question, that is, the conditions in which joint attention may emerge. Coordinated affective states, for example, are features of an interaction that are not immediately essential to attaining joint attention in both partners, but may be influential in creating jointness from which attention can be cultivated. Thus, the distinction between JA and JE lies in the level at which jointness between two partners is being studied.

It can be said, then, that joint engagement is a necessary condition for joint attention. Based on behavioral indicators of social relation used to asses jointness in children with autism, JA appears to encompass a variety of specific skills necessary for both attending to a person and object and spontaneously creating a shared point of reference with another person (Mundy et al., 2009). While many studies have investigated the significance of these skills in relation to outcomes in various developmental domains, fewer studies have examined the conditions in which these behaviors emerge. Even so, research on joint engagement to date appears to measure this phenomenon by relying primarily on the interpretative judgment of clinical experimenters, who use *global descriptions* derived from our understanding of joint attention skills used by infants and young children in interaction with their parents.

Global Measures of Joint Engagement

Indeed, several studies have developed behavioral coding schemes based on prior work on joint attention skills to assess JE in young children. Notably, Bakeman and Adamson's (1984) study that compared how infants socially orient themselves to their mothers and to their peers assessed jointness with various engagement states they entered in relation to people, objects, neither of the two, or both. Subsequent studies on joint engagement have gone on to adapt this coding scheme in assessing other variables with joint engagement in parents and children. Adamson, Bakeman, Deckner, and Romski (2009), for instance, expanded on their original coding scheme by using it within the framework of understanding how children attend to symbolic content. To this end, they conceptualized four "symbol-infused" engagement states to describe how variations in JE might predict expressive and receptive language outcomes. In a later study, Adamson, Bakeman, Deckner, and Nelson (2012) devised a 17-item battery, which included the original JE states and their symbol-infused variations, as well as three clusters of behaviors for the child, parent, and shared topic, in order to explore whether it was a viable method of capturing JE data in parent-child interactions. The authors found that ratings with this battery were ultimately able to capture a broad view of joint engagement in toddlers with TD and ASD.

Others like Miersschaut, Warreyn, and Roeyers (2011), who looked at whether mothers differed in their interaction styles when interacting with children with ASD and a sibling without ASD, used behavioral categories such as “social initiatives” and “social responses.” These categories encompassed a diverse number of actions, gestures, or verbal remarks from the mother. Social initiatives were differentiated into subcategories of declarative, imperative, and neutral behaviors, depending on whether mothers were sharing interest, requesting an action from the child, or neither of the two. Findings from the study indicated that mothers differed in their level of social responsiveness, but not in their initiatives towards their children with and without ASD, and alluded to a need for a better understanding of maternal strategies in families of children with ASD, especially in the context of family-based interventions.

In addressing this need, a more recent investigation by Mendive, Bornstein, and Sebastián (2013) integrated variables of JE from Bakeman and Adamson’s (1984) original study with maternal attention-directing strategies (ADS) used by Landry, Chapieski, and Schmidt (1986). The study hypothesized that specific maternal ADS (e.g., introducing, maintaining, redirecting), would precede specific infant engagement states (e.g., onlooking, supported joint, coordinated joint) and looked at mother-infant contingent relations among these strategies and engagement states. Their analysis revealed a significant role played by maternal attentional maintaining behaviors in facilitating infant coordinated joint engagement and a potential mediating role of maternal introducing and redirecting behaviors for socio-cognitive development in infants.

Toward Discrete Measures of Joint Engagement

While much of the existing literature on joint engagement in children and their parents has developed and adapted the use of coding schemes with global, interpretative measures, fewer studies have examined it with more discrete components of social behavior. In particular, these discrete measures have included rhythms of dialogue between two social partners (Feldstein, Konstantareas, Oxman, & Webster, 1982; Gratier et al., 2015; Northrup & Iverson, 2015), acoustic patterns of voice (Fusaroli, Lambrechts, Bang, Bowler, & Gaigg, 2017; Quigley, McNally, & Lawson, 2016), postural states in early development (Zapella et al., 2005), and proxemics (Abercrombie, 1971), the study of spatial relations between two individuals (Lindahl & Heimann, 1997; Pedersen, Livoir-Petersen, & Schelde, 1989).

Insights from Vocal Rhythm and Acoustics

Existing work on bi-directional coordination of speech patterns in parent-child dyads has revealed that this phenomenon can occur between mothers and their offspring in as early as infancy (Jaffe et al., 2001). In addition, this vocal coordination has been shown to be related to several aspects of development, such as language (Tamis-LeMonda, Bornstein, & Baumwell, 2001), as well as socio-cognitive awareness (Papoulidi, Papaeliou, & Samartzi, 2017). Given the links between joint attention and these same areas of early development, we can borrow insights from previous work on vocal rhythm and the human voice to capture patterns of joint engagement in children with ASD in more discrete ways.

Feldstein et al. (1982) investigated differences between the temporal synchrony of communication in youths with ASD with their parents and with experimenters. Results from the study indicated that youths with ASD demonstrated longer switching pauses—defined by McLaughlin (1984) as one’s silence between vocal turns—with their parents and both longer switching pauses and longer pauses with the experimenters. These results are indicative of the "asynchronies of autistic individuals in their kinesic responsiveness to others" (Feldstein et al., 1982, p. 452). Similarly, Northrup and Iverson (2015) compared vocal coordination in a group of high-risk infant siblings of children with ASD and their parents to a group of low-risk infants with their parents to predict later language outcomes. They found that, as a variable of vocal coordination, infant simultaneous speech (i.e, concurrent vocalizations of the infant with those of the parent) was predictive of later language delays (LD) in vocabulary development and communicative gestures. Dyads with LD infants were also less coordinated in their switching pauses.

Acoustic patterns of speech have also been shown to be a hopeful avenue for establishing directly measurable biomarkers of ASD in children. In a study on vocal prosody in ASD, Quigley et al. (2016) compared a group of infants at-risk for ASD with their mothers to a group of TD controls in their developmental trajectories of prosodic expression (i.e., changes in pitch and volume) between 12 and 18 months of age. They found that for dyads with at-risk infants, the trajectory of the development of prosodic expression was unexpected. In particular, at-risk infants increased in pitch and volume, contrary to expectations and results from previous studies on children with TD who exhibit a reduction in pitch and volume with age.

Recent work by Swanson et al. (2017) investigated the early language environment of 9-month-old infants at high risk for ASD using frequencies and durations as measures of vocalizations, conversational turns, and use of adult words. The authors found that, compared to infants at low risk (LR) for ASD, infants at high risk (HR) produced significantly more vocalizations—to the extent that they identified a subgroup of "hypervocal" infants who vocalized at a rate that was 2 *SD*s above the mean rate of low-risk infants. Compared to HR and

LR infants, these "hypervocal" infants also experienced more conversational turns and fewer adult words in interaction with their parents. In spite of these differences, "hypervocal" infants did not experience the same exaggerated increase in their conversational turns as they did in their vocalizations, which the authors attributed to excess vocalizations characterized by features associated with stereotypic behaviors. Findings from this study suggest that specific patterns in the rate of vocal behaviors seen in as early as infancy can lend valuable insight into early detection of autism spectrum disorder in children.

Insights from Proxemics and Early Motor Behaviors

Empirical studies in proxemics and early motor behaviors can also lend insight into devising more discrete ways of measuring joint engagement. Proximity is a very important component of communication, and spatial arrangements play an integral role in establishing and maintaining various types of social engagement (Abercrombie, 1971). For example, social proximity, defined by Lindahl and Heimann (1997) as any "pattern of behavior associated with the idea of 'relatedness to others'" (p. 84), has been shown to be a distinguishing feature of the way boys and girls jointly engage with their mothers. The results of the study revealed that mother-daughter dyads display significantly higher degrees of social proximity than mother-son dyads.

In the context of ASD, a number of studies have also confirmed the importance of proximity and sociomotor behaviors in approaches to early detection of the disorder, as well as in the quality of parent-child relationships. For example, Dissanayake and Crossley (1996) investigated whether children with ASD would direct proximity and social behaviors towards their mothers in two conditions—when alone together with their mothers and with their mothers and an unfamiliar female adult in the room. Comparing the children with ASD to children with

TD and Down syndrome, they found that, similar to children with TD and Down syndrome, the children with ASD exhibited proximity behaviors toward their mothers. Specifically, in the presence of an unfamiliar adult, the children with ASD increased their proximity behaviors toward their mothers. While their coding scheme included both global and discrete measures of proximity and social behaviors, including factors such as interpersonal distance and spatial orientation to smiling and mutual play, the authors ultimately provided meaningful evidence of attachment behaviors and the capacity to form person-specific bonds in children with ASD.

A more recent study by Zapella et al. (2005) retrospectively analyzed a series of home videos, recorded between birth and the age of 6 months, of boys who either were later diagnosed with ASD or had transient autistic symptoms until the age of three. Specifically, the study looked at general movements, in addition to eye contact, responsive smiling, and pre-speech vocalizations, of the different groups of infants. General movements were divided into writhing or "fidgety movements" across a variety of postural conditions depending on whether the infant was resting supine, hyperextending the trunk, or slumped, among other postural orientations. Most striking in their results was the high frequency of abnormal general movements in infants who were later diagnosed with ASD, particularly in their fast and exaggerated fidgety movements and monotonous writhing movements. The authors concluded that general movements might serve as a plausible early diagnostic marker of ASD in young children.

Purpose

The majority of studies on joint engagement in children with autism spectrum disorder and their families to date have used primarily global interpretations of social interaction. Taken together, however, these studies suggest the need for further research on measuring additional discrete elements of social interaction in joint engagement. The current study aims to contribute to the existing body of work on joint engagement in ASD by integrating specific variables from studies in proxemics, motor behaviors, and vocal rhythm. Specifically, the primary goal of the study was to devise a comprehensive parent-child coding system that can isolate **vocal and motor behaviors** as more discrete elements of joint engagement. To this end, the criterion validity of the coding system was assessed against established coding systems of joint engagement described in the review of the study by Mendive et al. (2013). Finally, differences in the frequencies and durations of dyadic vocal-motor behaviors between dyads of children with TD and ASD were tested.

Method

Participants

A total of *N* = 20 parent-child dyads were obtained from a larger study conducted at the University of Connecticut (PI: L. Naigles) that investigated longitudinal outcomes of expressive language development in young children with autism spectrum disorder (Tek, Mesite, Fein, & Naigles, 2014). Dyads were predominantly from White, middle-class suburban households. Across the dyads, 16 parents were mothers, 4 were fathers; 10 children were TD, and 10 were children with ASD. Children were between 3 and 4 years of age, and dyads were matched on child's age (*M* = 41.0 months, *SD* = 3.5, range = 35.0 to 48.8), and sex (10 boys, 10 girls).

Materials and Procedure

All participants' data were collected during home visits, where parents and children engaged in 30-minute semi-structured play sessions, the first half of which was adapted from the *Screening Tool for Autism in Two-Year-Olds* (STAT; Stone, Coonrod, & Ousley, 2000). Twelve separate play-based activities on the STAT engaged the child in pretend play with dolls, interactive play with blocks and nesting cups, imitative action play, requesting with snacks, and dyadic joint attention with a remote-controlled car, bag of toys, and balloon. All parents provided informed consent in the original study, which was approved by the Internal Review Boards of the University of Connecticut and Columbia University. Video-recorded play session data of parent-child dyads completing the joint-attention “Balloon Task” were obtained for the current study with permission. All videos of dyads with children who were out-of-frame of the video for over 50% of the total duration of the task were omitted from analyses. Videos of dyads who could not complete the “Balloon Task” according to given instructions due to the involvement of a sibling or a pet or ending the task prematurely were also omitted from analyses.

“Balloon Task.” Parents were handed a notecard by a research assistant in which instructions adapted from the STAT were provided for the “Balloon Task.” Each parent was instructed to hold up the deflated balloon, get the child’s attention verbally, blow up the balloon, hold it over their own head before alerting the child (e.g., “One... Two... Three!” or “Ready... Set... Go!”), and then release the inflated balloon so that it flies away. The task was repeated approximately three times per dyad, or until the research assistant verbally indicated for the parent to transition to the subsequent task.

Behavioral coding. We developed an original coding scheme, the Vocal-Motor Coding System (VMCS), with separate coding schemes for motoric and vocal behaviors based on existing studies on dyadic play and communication (Dissanayake & Crossley, 1996; Feldstein et al., 1982; Jaffe et al., 2001). Coding was based on recorded sessions of the “Balloon Task,” the durations of which ranged approximately from 2 to 5 min across dyads (*Mean* = 3.04 min, *SD* = 0.97). All dyadic interactions were coded by the author, who was blind to children’s ages and diagnoses.

Coding of motoric behaviors was performed in Datavyu (Datavyu Team, 2014), a video coding and data visualization tool for capturing behavioral data with times of onset and offset. Each behavioral occurrence was associated with times of onset and offset specific to the nearest hundredth millisecond, which were used to calculate the duration of each behavior. Frequencies and durations of these codes were quantified in an automated, custom program written in MATLAB (R2017a; The Mathworks, 2017) that counted the individual onsets of each behavioral event and then subtracted the time of onset from the time of offset for the duration of each behavioral event.

Vocal behaviors were "coded" in Audacity® (Audacity Team, 2017), a free, open-source audio editing and recording software, first by splitting each dyad's audio data into two tracks—one for the parent and one for the child—and then by reducing noise and silencing segments of no vocalization. Audio signals were subsequently passed through the Parent-Child Vocalization Analyzer, a custom program written in MATLAB, to quantify the frequencies and durations of audio codes of "0" for silence, "1" for the parent's vocalization, and "2" for the child's vocalization. Frequencies of loudness constitute the number of vocalization samples coded as "Loud" (see Figure A2) from the parents' and children's audio tracks, of which a majority had a sampling rate each of f_s = 32 kHz (or 32,000 samples per second). Durations, calculated in seconds, were derived from dividing the total frequency of loud vocalization samples by the sampling rate. Frequencies of parent and child vocalizations, pauses, and switching pauses indicate the number of occurrences of each behavior in the audio track of each social partner. Durations of vocal behaviors were computed as the number of samples between the onset and offset of a vocal behavior divided by the audio track's sampling rate.

Motoric behavioral coding scheme. The motor coding scheme (see Appendix) was developed to capture instances of seven possible physical dimensions of social interaction between a parent and their child, using a nominal scale to describe each of three dyadic dimensions: (1) Proximity; (2) Orientation; and (3) Object-Touch. *Proximity* was rated as either "Near" or "Far," depending on whether the child was observed to be within arm's reach of the parent. *Orientation* describes the alignments of the parent's and child's chest regions with respect to each other, and was rated as "Toward," "Parallel/Oblique," or "Away" (see Figure A1). *Object-Touch* describes whether the balloon was being touched or held by either the parent or child, and was rated as either "Yes" or "No" for the parent and child each.

Vocal behavioral coding scheme. The vocal coding scheme (see Appendix) encapsulates two dimensions in the vocal interaction of the dyad: (1) Loudness; and (2) Rhythmicity. *Loudness* was measured in MATLAB by the amplitude of a soundwave and described whether the amplitude of a vocalization was above or below a predetermined dyad-specific threshold. Thus, this dimension was rated nominally as either "Loud" or "Not Loud". *Rhythmicity* was adapted from the vocal rhythm variables described by Feldstein et al. (1982) and captured instances of parent and child vocalizations, pauses, and switching pauses within each interaction.

Criterion Measures

The criterion validity of our vocal-motor coding schemes was assessed against established coding systems for joint engagement discussed in work by Mendive et al. (2013). These existing coding systems were used to nominally rate each parent-child dyad on *Child Engagement State* (CES) and *Parental Attention-Directing Strategy* (PADS). Motoric behaviors were tested against six types of CES: (1) Unengaged; (2) Onlooking; (3) Object; (4) Person; (5) Supported Joint; and (6) Coordinated Joint. Vocal behaviors were assessed against five types of

PADS: (1) Introducing; (2) Maintaining; (3) Redirecting; (4) Independent Action; and (5) Not Responding. Frequencies and durations of Child Engagement States and Parental Attention-Directing Strategies observed in dyads were derived from coding the onset and offset of behaviors for each CES and PADS in Datavyu, which were then quantified in MATLAB.

Analytic Approach

Data analysis was performed in two parts. First, to assess the criterion validity of the Vocal-Motor Coding System, we calculated Spearman's rank correlation coefficients among the durations and frequencies of all 15 behavioral codes within the VMCS against the durations and frequencies of the six Child Engagement States and five Parental Attention-Directing Strategies referenced above. Additionally, we computed correlations between behaviors from the VMCS, conditioning on diagnosis (i.e., ASD and TD groups), to identify unique vocal-motor correlations within each respective group. Given that the frequency and duration data were non-normally distributed, nonparametric two-sample tests of differences in vocal-motor behaviors were assessed using Wilcoxon rank sum tests between groups. The assumption of equality of group variances was assessed using the Brown-Forsythe test. Vocal-motor behavioral variables that met the requisite assumptions were then examined further using the aforementioned non-parametric test of differences. Frequencies and durations of all 26 behavioral codes from the Vocal-Motor Coding System, as well as the coding schemes for Child Engagement States and Parental Attention-Directing Strategies (Mendive et al., 2013), are presented as *Median* (Interquartile Range). The significance level was established *a priori* as $\alpha = 0.05$.

Results

Descriptive Statistics

The median and interquartile ranges (IQR) of frequencies and durations of all behaviors (i.e., CES, PADS, motoric, and vocal) across the 20 dyads are given in Tables 1 and 2. Interquartile ranges of their distributions indicate the variability of the behavioral codes across all dyads. Frequencies of CES, PADS, and motoric behaviors indicate the number of instances in which they occurred during the "Balloon Task," which lasted roughly *3.03*(1.16) mins.

Descriptive data showed that the CES with the highest frequency across all dyads was Object engagement (*Mdn* = 3.00, *IQR* = 3.00), while the CES with the longest duration in an interaction was Coordinated Joint (*Mdn* = 16.90 sec, *IQR* = 18.79). Across all dyads, the PADS with the highest frequency was Maintaining (*Mdn* = 3.00, *IQR* = 1.75). Maintaining was also the strategy with the longest duration (*Mdn* = 36.93 sec, *IQR* = 28.27) among all five PADS.

Dyadic proximities coded as "Near" (i.e., an adult arm's length of distance between the parent and child) occurred *3.00*(4.50) times and lasted *51.59*(58.09) sec, while proximities coded as "Far" occurred approximately *2.00*(3.75) times and lasted *7.48*(13.17) sec between a parent and child during the "Balloon Task." Parents and children oriented themselves with the highest frequency in parallel or obliquely (*Mdn* = 6.50, *IQR* = 5.50) to each other during the task. Roughly, parents also held or touched the balloon *5.00*(2.25) times for *21.04*(8.12) sec, compared to their children who held or touched the balloon *4.00*(3.25) times for *11.19*(12.43) sec during the interaction.

Generally, parents vocalized *102.00*(29.00) times in an interaction, while children vocalized *64.00*(34.50) times. Parents paused *29.00*(19.25) times in an interaction, and children paused *11.00*(7.75) times. Switching pauses of parents (i.e., the periods of silence between a parent's vocalization and a child's vocal response) were observed approximately *15.00*(17.00) times and switching pauses of children (i.e., the periods of silence between a child's vocalization and a parent's vocal response), approximately *13.00*(15.00) times. Durations for parent and child vocalizations were *0.69*(0.04) and *0.61*(0.05) sec, respectively, while durations of parent and child pauses and switching pauses were *1.95*(1.16), *1.26*(0.86), *1.61*(1.31), and *1.50*(1.19) sec, respectively.

VMCS Criterion Validity

Spearman's rank-order correlations between dyads' vocal behaviors and PADS and between their motoric behaviors and CES are shown in Tables 2 and 3, respectively. Frequencies of parent pauses were found to be moderately and inversely correlated with introducing strategies ($\rho = -0.52$, $p < 0.05$), as well as with behaviors of non-responsiveness ($\rho = -0.57$, $p < 0.01$). Lengths of parent switching pauses showed a moderate positive monotonic relationship with the amount of time parents spent in behaviors of independent action ($\rho = 0.51$, $p < 0.05$). With child switching pauses, parental behaviors of redirecting were found to be moderately and inversely correlated ($\rho = -0.46$, $p < 0.05$) and parental behaviors of independent action were found to be moderately and positively correlated ($\rho = 0.65$, $p < 0.01$).

With respect to motoric behaviors and CES, there were moderate positive correlations between the frequencies of child unengaged behavior and proximity of far ($\rho = 0.45$, $p < 0.05$), as well as orientations facing away ($\rho = 0.47$, $p < 0.05$). The time children spent in unengaged behavior was also inversely correlated with the amount of time dyads spent oriented toward each other ($\rho = -0.58$, $p < 0.01$). The amount of time children were engaged with their parents was moderately correlated with time dyads spent oriented toward each other ($\rho = 0.61$, $p < 0.01$) and inversely correlated with the amount of time dyads spent oriented parallel or obliquely to each other ($\rho = -0.55$, $p < 0.05$). Lastly, moderate positive correlations were also found between instances of child engagement with an object in vicinity and instances of engagement with the balloon by the parent ($\rho = 0.54$, $p < 0.05$) and by the child ($\rho = 0.51$, $p < 0.05$).

Within Group Vocal-Motor Correlations

Spearman's rank-order correlation coefficients for frequencies of vocal rhythmic and motoric behaviors within dyads with ASD and TD are shown in Table 5. Correlation coefficients for durations of vocal and motoric behaviors are shown in Table 6. Within the ASD group, moderate-to-strong, positive relationships were found among frequencies of loud parent vocalizations and dyadic orientations of facing toward ($\rho = 0.66$, $p < 0.05$) and in parallel or oblique ($\rho = 0.63$, $p < 0.05$). A high volume of strong, positive correlations ($p < 0.001$) were also found between parent and child switching pause frequencies and proximities of near and far, as well as orientations in which dyads were facing toward and in parallel or obliquely to each other. Within the TD group, only the frequencies of parent pauses and instances in which parents held or touched the balloons showed a strong positive relationship ($\rho = 0.78$, $p < 0.01$).

Fewer correlations were found among durations of vocal and motoric behaviors within groups. There were moderate-to-strong, positive relationships between durations of parent pauses and time dyads with ASD spent near each other ($\rho = 0.81$, $p < 0.01$) and oriented in parallel or obliquely ($\rho = 0.68$, $p < 0.05$). Durations of loud vocalizations uttered by parents of children with ASD were also strongly and inversely correlated with the amount of time they spent near each other ($\rho = -0.90$, $p < 0.001$). Within the TD group, we found only durations of parent switching pauses to be significantly inversely correlated with the amount of time parents and children spent oriented in parallel or obliquely to each other ($\rho = -0.69$, $p < 0.05$).

Group Differences in Vocal and Motoric Behaviors

Dyads with ASD and TD did not differ significantly from each other in their vocal and motoric behaviors in terms of behavioral frequencies nor durations. Results from the Wilcoxon rank sum test of differences between the vocal and motoric behaviors in the ASD and TD groups are presented in Table 7.

Discussion

In the current study, we developed the Vocal-Motor Coding System as a novel approach to coding vocal and motoric behaviors of joint engagement in parent-child dyadic interactions. Historically, studies on joint engagement in children with autism spectrum disorder (ASD) have utilized primarily global measures of behavior derived from prior work on joint attention skills. Thus, the main goal of the current study was to contribute to existing work on JE in dyads with ASD by integrating more discrete measures of behavior in assessing "jointness" (Hobson & Hobson, 2011). Specifically, we integrated motoric and vocal behavioral variables based on the work of prior studies of proxemics, motor behaviors, and vocal rhythm into our novel coding system.

The VMCS as a Viable Instrument

Results from our analyses of the criterion validity of the VMCS and of differences in vocal-motor behaviors between typically developing dyads and dyads with ASD suggest moderate correlations between the VMCS and specific aspects of the coding schemes for joint engagement and parental strategies established in work by Mendive et al. (2013).

Correlations between vocal behaviors and PADS. Our analysis revealed moderate correlations between vocal behaviors and Parental Attention-Directing Strategies (PADS). Specifically, inverse correlations were found between the frequency of parent pauses and two types of PADS—introducing and not responding. This finding suggests that the *Rhythmicity* dimension of parent pauses can adequately describe periods of parental non-responsiveness towards the child and introducing strategies. A moderate positive correlation was also found between the amount of time parents spent not attending to their children's behaviors and the lengths of parents' and children's switching pauses. Since switching pauses signify the periods of silence between the vocal turns of the parent and child, this positive correlation appears to suggest that our vocal variable of switching pauses can indicate whether parents and children are jointly engaged, particularly on the condition that they are engaged in separate activities. Indirect support for these correlations can be derived from prior work demonstrating the efficacy of relationship-focused interventions that emphasize the use of more responsive interactive strategies in mothers to improve the sociable behaviors of children with ASD (Mahoney & Perales, 2003). In summary, findings from the current study offer meaningful contributions to the design of such parent-mediated interventions that may utilize vocal variables such as pauses and switching pauses to measure levels of parental responsiveness.

Correlations between motoric behaviors and CES. Our analysis of the frequencies and durations of motoric behaviors and Child Engagement States (CES) also showed moderate correlations that support the viability of the VMCS. In particular, how often children were unengaged in the activity was inversely correlated with how long parents and children spent oriented toward each other and was positively correlated with how often parents and children were far in proximity or oriented away from each other. These correlations suggest that our motoric coding scheme may assess whether a parent-child dyad is mutually engaged depending on whether the parent and child are positioned far (i.e., out of arm's reach) or oriented away from each other. Moderate correlations between the amount of time children spent engaged with only their parents and how long parents and children spent oriented toward or in parallel or obliquely to each other during the activity also suggest the *Proximity* and *Orientation* codes may be used to assess whether the child is engaged with the parent. These implications of the motoric dimensions of *Proximity* and *Orientation* in the VMCS can be linked to previous findings from studies investigating proximity and orientation as measures of child engagement. For example, Doussard-Roosevelt, Joe, Bazhenova, and Porges (2003) demonstrated that children with ASD showed greater responsiveness to maternal approaches involving increased physical proximity. Pedersen et al. (1989) also demonstrated that "autistic children with full syndrome" (classified according to the *Diagnostic and Statistical Manual of Mental Disorders* [3rd ed.; American Psychiatric Association, 1980]), engaged more with adults within shorter interpersonal distances. Additionally, Dissanayake and Crossley (1996) showed that, as evidence of attachment behaviors, children with ASD also appear to engage with their mothers through measures of frontal body orientation and proximity. In contrast to these studies, use of the combination of both dyadic proximity and orientation in our coding system may serve as an advantage in refining our understanding of how levels of social engagement between parents and children may be impacted by the way dyads physically position and orient themselves in space.

Finally, we found moderate positive correlations between child engagement in object play and touching of the balloon by the parent and child, suggesting that the motoric dimension of *Object-Touch* may point toward whether a child is engaged in object play by him or herself depending on the persistence with which the parent and child handle the balloon separately. Given the pooling of data from both the ASD and TD groups in these correlations, a meaningful conclusion we can draw at present is that our dimension of *Object-Touch* may be able to capture instances of the CES of Object engagement. Doussard-Roosevelt et al. (2003) have also shown that children with ASD display increased responsiveness to maternal approach behaviors involving nonverbal object use. To support and elucidate this claim, further analyses are needed to parse through these behaviors and determine differences in ways children with TD and ASD might disengage from the joint activity in response to parental handling of the balloon or object of interest.

Vocal-Motor Strategies for Joint Engagement

While the median values appeared to be descriptively different between the behavioral frequencies and durations of the two groups, the high variance in our measures of vocal and motoric behaviors in the ASD group may have contributed to a lack of significant differences between dyads with TD and ASD. The wide dispersion of vocal-motor measures seen in our ASD group may be attributed to the phenotypic heterogeneity of the disorder in children (Kim, Macari, Koller, & Chawarska, 2015). Nonetheless, we found that a majority of the vocal and motoric behaviors of dyads with ASD were strongly and positively correlated with each other. Specifically, loud parental vocalizations were related to dyadic proximity and whether children with ASD and their parents were oriented toward and in parallel or obliquely to each other. Pauses of parents and switching pauses of both partners were also strongly and positively related to their dyadic proximity and orientation. Taken together, these correlations reveal interrelationships between vocal-motor behaviors and parental vocal strategies that may be unique to joint engagement in dyads with ASD. Notably, parents of children with ASD were likely to spend less time vocalizing loudly the more time they spent in close proximity with their children. This finding may be a strong indicator of a level of attunement and responsiveness in parents of children with ASD, as supported by prior work on parental synchronization of behaviors to children within parent-child dyads with ASD (Siller & Sigman, 2002). The positive correlation between parental pauses and dyadic proximity also suggests that a parent likely paused in their vocalizations in either state of dyadic proximity. This frequency of pauses by the parent may be due to the infrequency of vocalizations in response to the parent by the child with ASD, as evidenced by existing literature on the social responsiveness of children with ASD (Dawson & Adams, 1984; Mundy, Sigman, Ungerer, & Sherman, 1986b). Conversely, dyads with TD children showed a positive association between rates of parental engagement with the balloon and pauses in parental vocalizations, suggesting that a relatively high degree of silence from the parent occurred with parental handling of the balloon. Consistent with findings from previous studies comparing the levels of joint engagement between dyads with ASD and TD (Adamson et al., 2012), this outcome may be an implicit indication of the fact that parents and their TD children are more likely to be in coordinated joint engagement, allowing for fewer parental vocalizations and more periods of silence between both partners.

Parental attunement. Given the positive associations between dyadic proximity and vocal behaviors, such as parental loudness, pausing, and turn-taking as observed from children's switching pauses, vocal-motor behaviors may serve as a quantifiable measure of parental attunement in the parent-child dyad with ASD. Attunement is defined as "a parent's ability to understand their child, act on that understanding and adjust to their child's needs" (Zand et al., 2014, p. 7). Through the frequency and duration of pauses, as well as the frequency of their vocal turn-taking behaviors, parents of children with ASD may be demonstrating specific vocal-motor behaviors indicative of the quality of their attunement, while striving to achieve jointness with their children. Although no significant correlations were found between these variables within dyads with typically developing children, results from our test of differences precludes us from making a meaningful comparison between the groups' parental strategies. However, median values of the frequency of these vocal behaviors all appear to be greater in the ASD group than in the TD group, which may suggest increased use of strategies by parents of children with ASD that might become increasingly apparent in a larger cohort. These initial findings may support results from prior studies on increased use of parental strategies, such as psychological control (Gau et al., 2010) and power assertion (Riany, Cuskelly, & Meredith, 2017), that can be used in the context of assessing parent-led interventions for improving joint engagement outcomes in parent-child dyads with ASD.

Limitations of the Study

A few limitations to the current study deserve discussion. The first limitation is the inconsistency of the naturalistic setting in which play sessions were carried out and recorded. Research in proxemics has referenced the effects of furniture arrangements on our perception and use of proximity and distance in social interaction (Abercrombie, 1971). In the case of the video-recorded parent-child play sessions used in this study, the placement and availability of furniture and competing toys in the room were a contributing factor to the challenge some dyads faced to complete the "Balloon Task." For instance, the accessibility of an open sofa in the living room invited opportunities for the child to roam away from the center of the joint task and climb onto the sofa, distracted and unengaged or intending to watch his or her parent from afar. Similarly, competing toys in the vicinity of the dyad might have also negatively impacted the motivation of the child to engage in the joint-attention task. Setting guidelines for the naturalistic environment prior to the visit would offer a more streamlined assessment and remove possible bias in the scoring of each dyad.

A second limitation of the study is the scarcity of fathers in the sample ($n = 4$). Due to the small number of fathers within our dyads, we could not assess the effect of parent gender on the vocal-motor behaviors observed in each dyadic interaction. Whether fathers demonstrate more "near" behaviors with their children or mothers utter more vocalizations towards their children, for example, could not be determined. Moreover, the variability in parent gender in our sample could have diluted observable differences in behaviors between our dyad groups.

A third limitation of the present study is the familiarity of our parents and children with the specific nature of the "Balloon Task." The majority of video-recorded play sessions took place during the fifth or sixth visit (separated by 4 months) by the experimenters from the original study (Tek et al., 2014). Barring cultural differences, the joint-attention activity with the balloon was observed to be more or less a familiar activity to most of the typically developing children by the fifth and sixth visits. Whether familiarity with the activity or use of a balloon in general contributed to the ease in completion of the task for some of the dyads eludes the parameters of this study. Obtaining a history of families' experiences using balloons with a brief and informal questionnaire, for example, would lend valuable insight into the effect of dyads' familiarity with the nature of the task on their observable behavioral outcomes.

A final limitation is the lack of scores of children's neuropsychological functions. Information such as severity ratings of their ASD diagnoses and cognitive and language test scores could ultimately serve as additional covariates that may help to explain some of the high variance found in the behaviors of our dyads with ASD.

Finally, our study was delimited to a small number of design factors that, if revised in future work, may allow for more specific research questions around the interrelationships of vocal-motor behaviors in parent-child dyadic interactions. One of these delimitations was the specific parent-child activity chosen to help achieve our research aims. While the "Balloon Task" was a highly physically engaging activity, and thus served as an ideal context for observing motoric behaviors, the balloon itself was not a highly motivating object to the child. Two other JA activities among the 12 play-based activities adapted from the STAT (Stone et al., 2000) involved the use of a remote-controlled car and a bag of toys. None of the children in our dyads who attempted to inflate the balloon on his or her own was able to successfully do it. This evidently skewed agency over the joint activity to the parent, leaving the child as a passive observer, which may have impacted the child's level of motivation to spontaneously initiate joint attention. A final delimitation was the analytic approach we used in investigating interrelationships among the vocal-motor variables of our dyads. While frequencies and durations can capture a high degree of information about behaviors, they may not be adequate in explaining how certain behaviors between coding schemes and between social partners related to each other; rather, sequential analysis of behaviors, as discussed subsequently, may deliver a more sophisticated understanding of the temporal associations between vocal and motoric behaviors within parent-child dyads.

Strengths

Despite the noted limitations, our study is one of the first studies to develop a behavioral coding system for joint engagement that integrates a number of variables from prior work in proxemics, motor behaviors, and vocal rhythm. Although previous studies have used discrete elements of social interaction, such as interpersonal distance, orientation, and object-engagement, to measure joint engagement in parents and children, few have used a coding strategy that encompasses all of these variables simultaneously. The breadth of our proposed coding system offers an opportunity to capture joint engagement in more discrete ways.

Future Directions for Research

Findings from the present study implicate a number of directions for future research in the area of joint engagement in children with ASD and their parents. The goal of the current study was to develop the Vocal-Motor Coding System as a new and viable approach to evaluating joint engagement in parents and children with ASD with more discrete and quantifiable measures. Future work should build upon the findings of our study by assessing a number of hypotheses related to the specificity of the VMCS and the interaction effects of variables such as parent and child gender or linguistic aspects of communication on vocal-motor behaviors of joint engagement.

Use of the VMCS in future work on contingencies between vocal and motor behaviors of dyads in interaction would likely lend additional credibility to our system. Time-series analysis of these behaviors would describe how some vocal and motor behaviors are temporally related within and between the social partners (Bakeman & Gottman, 1997). Additionally, this type of analysis would address questions related to specific patterns in vocal-motor strategies used by a certain dyad or social partner.

We also recommend future studies to investigate the effects of gender on vocal-motor behavioral outcomes of joint engagement. Specifically, how do mothers and fathers differ between their vocal-motor interactive styles with their children? How do boys and girls differ in interacting with their parents? These questions should be addressed in order to further investigate parental strategies and approach behaviors in the context of parent-mediated interventions within families of children with ASD. There remains a need in the current body of literature to assess the effectiveness of parent-led interventions designed to optimize social interactions between parents and children with ASD (Chiang et al., 2016; Patterson et al., 2014;

Shire, Gulsrud, & Kasari, 2016). Future work can begin to address this need by integrating our novel vocal-motor coding system as a possible outcome measure for change.

References

Abercrombie, M. L. J. (1971). Face to face—Proximity and distance. *Journal of Psychosomatic Research, 15*, 395–402.

Adamson, L. B., Bakeman, R., Deckner, D. F., & Nelson, B. P. (2012). Rating parent-child interactions: Joint engagement, communication dynamics, and shared topics in autism, Down syndrome, and typical development. *Journal of Autism and Developmental Disorders, 42*, 2622–2635.

Adamson, L. B., Bakeman, R., Deckner, D. F., & Romski, M. (2009). Joint engagement and the emergence of language in children with autism and Down syndrome. *Journal of Autism and Developmental Disorders, 39*(1), 84–96.

Adamson, L. B., Bakeman, R., & Suma, K. (2017). An expanded view of joint attention: Skill, engagement, and language in typical development and autism. *Child Development, 00*(0), 1–18.

American Psychiatric Association. (1980). *Diagnostic and Statistical Manual of Mental Disorders* (3rd ed.). Washington, DC: Author.

Audacity Team. (2017). Audacity(R): Free audio editor and recorder [Computer application] (Version 2.1.3). Retrieved from http://www.audacityteam.org

Bakeman, R., & Adamson, L. B. (1984). Coordinating attention to people and objects in mother-infant and peer-infant interaction. *Child Development, 55*(4), 1278–1289.

Bakeman, R., & Gottman, J. M. (1997). *Observing interaction: An introduction to sequential anaysis* (2nd ed.). Cambridge, UK: Cambridge University Press.

Bottema-Beutel, K., Yoder, P. J., Hochman, J. M., & Watson, L. R. (2014). The role of supported joint engagement and parent utterances in language and social communication development in children with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 44*(9), 2162–2174.

Charman, T. (2003). Why is joint attention a pivotal skill in autism? *Philosophical Transactions of the Royal Society B: Biological Sciences, 358*(1430), 315–324.

Chiang, C.-H., Chu, C.-L., & Tsung-Chin Lee, T.-C. (2016). Efficacy of caregiver-mediated joint engagement intervention for young children with autism spectrum disorders. *Autism, 20*(2), 172–182.

Datavyu Team. (2014). Datavyu: A video coding tool (Version 1.3.4). New York, NY: Databrary Project, New York University. Retrieved from http://datavyu.org

Dawson, G., & Adams, A. (1984). Imitation and social responsiveness in autistic children. *Journal of Abnormal Child Psychology, 12*, 209–225.

Dawson, G., Toth, K., Abbott, R., Osterline, J., Munson, J., Estes, A., & Liaw, J. (2004). Early social attention impairments in autism: Social orienting, joint attention, and attention to distress. *Developmental Psychology, 40*(2), 271–283.

De Barbaro, K., Johnson, C. J., Forster, D., & Deák, G. O. (2016). Sensorimotor decoupling contributes to triadic attention: A longitudinal investigation of mother-infant-object interactions. *Child Development, 87*(2), 494–512.

Dissanayake, C., & Crossley, S. A. (1996). Proximity and sociable behaviors in autism: Evidence for attachment. *Journal of Child Psychology and Psychiatry, 37*(2), 149-156.

Doussard-Roosevelt, J., Joe, C. M., Bazhenova, O. V., & Porges, S. W. (2003). Mother-child interaction in autistic and nonautistic children: Characteristics of maternal approach behaviors and child social responses. *Development and Psychopathology, 15*, 277-295.

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*, 451–460.

Fusaroli, R., Lambrechts, A., Bang, D., Bowler, D. M., & Gaigg, S. B. (2017). Is voice a marker for autism spectrum disorder? A systematic review and meta-analysis. *Autism Research, 10*, 384–407.

Gau, S. S., Chou, M. C., Lee, J. C., Wong, C. C., Chou, W. J., Chen, M. F., . . . Wu, Y. Y. (2010). Behavioral problems and parenting style among Taiwanese children with autism and their siblings. *Psychiatry and Clinical Neuroscience, 64*(1), 70-78.

Girolametto, L., Verbey, M., & Tannock, R. (1994). Improving joint engagement in parent-child interaction: An intervention study. *Journal of Early Intervention, 18*(2), 155–167.

Gratier, M., Devouche, E., Guellai, B., Infanti, R., Yilmaz, E., & Parlato-Oliveira, E. (2015). Early development of turn-taking in vocal interaction beween mothers and infants. *Frontiers in Psychology, 6*, 236–245.

Gulsrud, A. C., Jahromi, L. B., & Kasari, C. (2010). The co-regulation of emotions between mothers and their children with autism. *Journal of Autism and Developmental Disorders, 40*(2), 277–237.

Hobson, P., & Hobson, J. (2011). Joint attention or joint engagement? Insights from autism. In A. Seemann (Ed.), *Joint attention: New developments in psychology, philosophy of mind, and social neuroscience* (pp. 115–135). Cambridge, MA: MIT Press.

Jaffe, J., Beebe, B., Feldstein, S., Crown, C. L., Jasnow, M. D., Rochat, P., & Stern, D. N. (2001). Rhythms of dialogue in infancy: Coordinated timing in development. *Monographs of the Society for Research in Child Development*, i–149.

Kim, S. H., Macari, S., Koller, J., & Chawarska, K. (2015). Examining the phenotypic heterogeneity of early autism spectrum disorder: Subtypes and short-term outcomes. *Journal of Child Psychology and Psychiatry, 57*(1), 93–102.

Landry, S. H., Chapieski, M., & Schmidt, M. (1986). Effects of maternal attention-directing strategies on preterms' response to toys. *Infant Behavior and Development, 9*, 257–269.

Lindahl, L. B., & Heimann, M. (1997). Research report: Social proximity in early mother-infant interactions: Implications for gender differences? *Early Development and Parenting, 6*(2), 83–88.

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Jr., 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, 30*(3), 205–223.

Mahoney, G., & Perales, F. (2003). Using relationship-focused intervention to enhance the social–emotional functioning of young children with autism spectrum disorders. *Topics in Early Childhood Special Education, 23*(2), 74–86.

McLaughlin, M. L. (1984). *Conversation: How talk is organized*. Beverly Hills: Sage Publications.

Mendive, S., Bornstein, M. C., & Sebastián, C. (2013). The role of maternal attention-directing strategies in 9-month-old infants attaining joint engagement. *Infant Behavior and Development, 36*, 115–123.

Miersschaut, M., Warreyn, P., & Roeyers, H. (2011). What is the impact of autism on mother-child interactions within families with a child with autism spectrum disorder? *Autism Research, 4*(5), 358–367.

Mundy, P., & Newell, L. (2007). Attention, joint attention, and social cognition. *Current Directions in Psychological Science, 16*(5), 269–274.

Mundy, P., Sigman, M., Ungerer, J., & Sherman, T. (1986a). Defining the social deficits of autism—The contribution of nonverbal communication measures. *Journal of Child Psychology and Psychiatry, 27*, 657–669.

Mundy, P., Sigman, M., Ungerer, J., & Sherman, T. (1986b). Defining the social deficits of autism: The contribution of nonverbal communication measures. *Journal of Child Psychology and Psychiatry, 27*, 657–669.

Mundy, P., Sullivan, L., & Mastergeorge, A. M. (2009). A parallel and distributed-processing model of joint attention, social cognition and autism. *Autism Research, 2*(1), 2–21.

Naber, F. B. A., Swinkels, S. H. N., Buitelaar, J. K., Dietz, C., Van Daalen, E., Bakermans-Kranenburg, M. J., . . . Van Engeland, H. (2007). Joint attention and attachment in toddlers with autism. *Journal of Abnormal Child Psychology, 35*(6), 899–911.

Northrup, J. B., & Iverson, J. M. (2015). Vocal coordination during early parent-infant interactions predicts language outcome in infant siblings of children with autism spectrum disorder. *Infancy, 20*(5), 523–547.

Papoulidi, A., Papaeliou, C. F., & Samartzi, S. (2017). Rhythm in interactions between children with autism spectrum disorder and their mothers. *Timing & Time Perception, 5*(1), 5–34.

Patterson, S. Y., Elder, L., Gulsrud, A., & Kasari, C. (2014). The association between parental interaction style and children's joint engagement in families with toddlers with autism. *Autism, 18*(5), 511–518.

Pedersen, J., Livoir-Petersen, M. F., & Schelde, J. T. M. (1989). An ethological approach to autism: An analysis of visual behaviour and interpersonal contact in a child versus adult interaction. *Acta Psychiatrica Scandinavica, 80*(4), 346–355.

Quigley, J., McNally, S., & Lawson, S. (2016). Prosodic patterns in interaction of low-risk and at-risk-of-autism spectrum disorders infants and their mothers at 12 and 18 months. *Language Learning and Development, 12*(3), 295–310.

Riany, Y. E., Cuskelly, M., & Meredith, P. (2017). Parenting Style and Parent–Child Relationship: A Comparative Study of Indonesian Parents of Children with and without Autism Spectrum Disorder (ASD). *Journal of Child and Family Studies, 26*(12), 3559-3571.

Shire, S. Y., Gulsrud, A., & Kasari, C. (2016). Increasing responsive parent-child interactions and joint engagement: Comparing the influence of parent-mediated intervention and parent psychoeducation. *Journal of Autism and Developmental Disorders, 46*(5), 1737-1747.

Sigman, M., Dijamco, A., Gratier, M., & Rozga, A. (2004). Early detection of core deficits in autism. *Journal of Mental Retardation and Developmental Disabilities Research Reviews, 10*, 221–223.

Sigman, M., Mundy, P., Sherman, T., & Ungerer, T. (1986). Social interactions of autistic, mentally retarded, and normal children and their caregivers. *Journal of Child Psychology and Psychiatry, 27*(5), 647–656.

Siller, M., & Sigman, M. (2002). The behaviors of parents of children with autism predict the subsequent development of their children’s communication. *Journal of Autism and Developmental Disorders, 32*(2), 77–89.

Siller, M., Swanson, M., Gerber, A., Hutman, T., & Sigman, M. (2014). A parent-mediated intervention that targets responsive parental behaviors increases attachment behaviors in children with ASD: Results from a randomized clinical trial. *Journal of Autism and Developmental Disorders, 44*(7), 1720-1732.

Stone, W. L., Coonrod, E. E., & Ousley, O. Y. (2000). Brief report: Screening Tool for Autism in Two-Year-Olds (STAT): Development and preliminary data. *Journal of Autism and Developmental Disorders, 30*(6), 607–612.

Swanson, M. R., Shen, M. D., Boyd, B., Elison, J. T., Parish-Morris, J., Botteron, K., . . . Piven, J. (2017). Naturalistic language recordings reveal “hypervocal” infants at high familial risk for autism. *Child Development*, 1–14.

Tamis-LeMonda, C. S., Bornstein, M. H., & Baumwell, L. (2001). Maternal responsiveness and children’s achievement of language milestones. *Child Development, 72*(3), 748–767.

Tek, S., Mesite, L., Fein, D., & Naigles, L. (2014). Longitudinal analyses of expressive language development reveal two distinct language profiles among young children with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 44*(1), 75–89.

The Mathworks. (2017). MATLAB R2017a [Computer software]. Natick, MA: The Mathworks, Inc.

Ting, V., & Weiss, J. A. (2017). Emotion regulation and parent co-regulation in children with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 47*(3), 680–689.

Tomasello, M. (1995). Joint attention as social cognition. In C. Moore & P. J. Dunham (Eds.), *Joint attention; its origins and role in development* (pp. 103–130). Hillsdale, NJ: Erlbaum.

Zand, D., Pierce, K., Thomson, N., Baig, M. W., Teodorescu, C., Nibras, S., & Maxim, R. (2014). Social competence in infants and toddlers with special healthcare needs: The roles of parental knowledge, expectations, atttunement, and attitudes toward child independence. *Children (Basel), 1*(1), 5-20.

Zapella, M., Enspieler, C., Bartl-Pokorny, K. D., Krieber, M., Coleman, M., Bölte, S., & Marschik, P. B. (2005). What do home videos tell us about early motor and socio-communicative behaviours in children with autistic features during the second year of life—An exploratory study. *Early Human Development, 91*, 569–575.

Zhou, T., & Yi, C. (2014). Parenting styles and parents' perspectives on how their own emotions affect the functioning of children with autism spectrum disorders. *Family Process, 53*, 67–79.

Zukow-Goldring, P., & Arbib, M. A. (2007). Affordances, effectivities, and assisted imitation: Caregivers and the directing of attention. *Neurocomputing, 70*(13-15), 2181–2193.

Table 1

Descriptive Statistics of Criterion Measures by Frequency and Duration

	Frequency (counts)		***Duration (sec)***	
Behavior	Median	IQR	Median	IQR
Child Engagement States				
Unengaged	1.50	3.00	3.61	7.61
Onlooking	3.00	1.50	10.30	7.66
Object	3.00	3.00	8.82	4.99
Person	1.00	2.00	3.50	11.01
Supported Joint	1.00	2.00	8.57	16.07
Coordinated Joint	2.50	3.25	16.90	18.79
Parental Attention-Directing Strategies				
Introducing	1.00	1.00	4.69	7.17
Maintaining	3.00	1.25	36.93	28.27
Redirecting	2.00	2.00	6.32	5.09
Independent Action	1.00	1.50	6.33	5.13
Not Responding	0.00	1.00	0.00	6.91

Note. Durations denote the amount of time spent in seconds. Frequencies denote the number of coded occurrences of each Child Engagement State and Parental Attention-Directing Strategy across all dyads.

Table 2

Descriptive Statistics of Vocal and Motor Behaviors by Frequency and Duration

	Frequency (counts)		***Duration (sec)***	
Behavior	Median	IQR	Median	IQR
Loudness				
Parent	4551.50	7539.00	0.14	0.24
Child	2465.50	4701.50	0.08	0.15
Vocalizations				
Parent	102.00	29.00	0.69	0.04
Child	64.00	34.50	0.61	0.05
Pauses				
Parent	29.00	19.25	1.95	1.16
Child	11.00	7.75	1.26	0.86
Switching Pauses				
Parent	15.00	17.00	1.61	1.42
Child	13.00	15.00	1.50	1.19
Proximity				
Near	3.00	4.50	51.59	58.09
Far	2.00	3.75	7.48	13.17
Orientation				
Toward	3.00	5.25	7.80	14.20
Parallel/Oblique	6.50	5.50	18.74	29.27
Away	3.00	3.50	4.26	4.17
Object–Touch				
Parent–Balloon	5.00	2.25	21.04	8.12
Child–Balloon	4.00	3.25	11.19	12.43

Note. Frequencies of the Loudness dimension indicate total counts of vocalization samples that were coded as “Loud.” Frequencies of vocal rhythm variables indicate number of occurrences of vocalizations, pauses, and switching pauses.

Table 3

Spearman Correlation Coefficients of Vocal Behaviors and Parental Attention-Directing Strategies (PADS)

	Introducing	Maintaining	Redirecting	Independent Action	Not Responding
Parent Loudness					
Frequency	−0.06	0.34	−0.04	0.36	−0.33
Duration	−0.26	0.11	−0.04	0.03	−0.37
Child Loudness					
Frequency	−0.21	0.18	−0.11	0.13	−0.37
Duration	−0.36	0.25	−0.10	−0.03	−0.36
Parent Vocalizations					
Frequency	0.07	−0.03	−0.16	−0.06	−0.26
Duration	0.28	0.10	0.13	0.04	0.28
Child Vocalizations					
Frequency	0.27	−0.09	−0.15	−0.14	−0.24
Duration	−0.03	0.09	0.37	0.12	−0.14
Parent Pauses					
Frequency	-0.52^{*}	0.08	0.18	0.19	-0.57^{**}
Duration	−0.06	−0.17	−0.19	0.19	0.41
Child Pauses					
Frequency	0.23	0.15	0.21	0.16	0.03
Duration	0.14	−0.30	0.09	0.25	0.34
Parent Switching Pauses					
Frequency	0.06	0.13	0.13	−0.01	−0.35
Duration	−0.04	−0.34	−0.24	0.51^{*}	0.43
Child Switching Pauses					
Frequency	−0.11	0.16	0.24	0.04	−0.36
Duration	−0.34	0.05	-0.46^{*}	0.65^{**}	0.28

Significance: * $p < 0.05$. ** $p < 0.01$.

Table 4

Spearman Correlation Coefficients of Motor Behaviors and Child Engagement States (CES)

	Unengaged	Onlooking	Object	Person	Supported Joint	Coordinated Joint
Near						
Frequency	0.40	0.35	0.27	0.19	0.22	−0.03
Duration	−.16	.22	.18	−.15	.15	.27
Far						
Frequency	0.45*	0.31	0.25	0.28	0.07	0.04
Duration	0.20	−0.22	−0.20	0.20	−0.28	.09
Toward						
Frequency	0.22	0.17	0.17	0.40	0.15	0.08
Duration	−0.58**	−0.39	0.35	0.61**	−0.39	0.36
Parallel/Oblique						
Frequency	0.37	0.16	0.19	0.30	0.24	0.03
Duration	0.05	0.41	0.12	−0.55*	0.01	0.06
Away						
Frequency	0.47*	0.39	0.34	0.12	0.33	0.08
Duration	0.04	−0.08	−0.07	0.21	−0.38	0.29
Parent–Balloon						
Frequency	−0.25	0.06	0.54*	−0.01	0.34	0.18
Duration	0.08	−0.10	0.22	0.03	−0.02	0.08
Child–Balloon						
Frequency	−0.03	0.17	0.51*	0.16	0.42	0.27
Duration	−0.18	0.08	0.50*	0.17	−0.01	0.42

Significance: * $p < 0.05$. ** $p < 0.01$.

Table 5

Spearman Correlation Coefficients of Vocal Rhythmic and Motor Behavioral Frequencies of Dyads with ASD and TD

	Motoric Behaviors													
	Near		*Far*		*Toward*		*Parallel/Oblique*		*Away*		*Parent–Balloon*		*Child–Balloon*	
Rhythmicity	ASD	TD	ASD	TD	ASD	TD	ASD	TD	ASD	TD	ASD	TD	ASD	TD
Vocalization														
Parent	0.10	0.17	0.03	0.52	–0.06	0.58	–0.10	0.37	–0.02	0.27	–0.40	–0.26	–0.01	0.12
Child	0.21	0.15	0.10	0.55	0.11	0.36	0.00	0.18	–0.01	0.42	–0.27	–0.54	0.04	–0.12
Pauses														
Parent	0.66*	0.44	0.65*	0.21	0.58	0.40	0.60	0.39	0.57	–0.18	0.26	0.78**	0.47	0.31
Child	0.44	0.10	0.37	0.11	0.58	–0.33	0.48	–0.14	0.45	0.40	0.77**	–0.26	0.54	–0.46
Switching Pauses														
Parent	0.89***	–0.13	0.85**	0.11	0.90***	0.49	0.84**	0.52	0.69*	0.28	0.48	0.04	0.78**	0.27
Child	0.93***	–0.34	0.90***	–0.10	0.93***	0.27	0.95***	0.28	0.87**	0.00	0.54	0.21	0.88***	0.31

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

Table 6

Spearman Correlation Coefficients of Vocal and Motor Behavioral Durations of Dyads with ASD and TD

	Motoric Behaviors													
	Near		*Far*		*Toward*		*Parallel/Oblique*		*Away*		*Parent–Balloon*		*Child–Balloon*	
Vocal Behaviors	ASD	TD	ASD	TD	ASD	TD	ASD	TD	ASD	TD	ASD	TD	ASD	TD
Loudness														
Parent	–0.90***	–0.20	0.51	0.33	0.45	0.23	–0.78	–0.20	0.18	0.43	0.26	–0.02	0.05	0.12
Child	–0.55	–0.48	0.45	0.53	0.58	0.21	–0.52	–0.52	0.19	0.16	–0.37	–0.03	0.13	–0.39
Vocalization														
Parent	0.24	0.32	–0.41	0.01	–0.17	-0.10	–0.05	0.25	-0.04	0.26	0.39	0.18	0.04	0.41
Child	–0.39	0.19	0.23	0.47	–0.31	-0.04	–0.26	0.26	-0.01	0.50	0.14	–0.08	0.09	0.52
Pauses														
Parent	0.81**	0.27	–0.45	–0.52	–0.05	–0.18	0.68*	0.44	–0.22	–0.25	–0.35	0.12	0.33	0.15
Child	0.14	0.31	0.33	0.33	0.17	–0.04	–0.09	–0.22	0.64	0.12	–0.41	–0.41	0.31	0.27
Switching Pauses														
Parent	0.48	0.04	0.23	–0.27	0.27	–0.69*	0.36	0.73	0.48	–0.45	–0.41	0.13	0.59	–0.07
Child	0.44	–0.18	–0.18	–0.28	0.60	–0.58	0.14	0.45	0.17	–0.55	–0.43	0.41	0.63*	–0.45

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

Table 7

Comparison of ASD and TD Dyads in Vocal and Motoric Behaviors

	ASD		TD		
Variables	*Median*	*IQR*	*Median*	*IQR*	p^a
Vocal Frequencies[b]					
Parent Loudness	9756.00	9194.00	3808.00	1893.00	0.21
Child Loudness	2836.00	5117.50	1442.50	1959.75	0.73
Parent Pauses	29.00	17.50	29.00	17.00	0.85
Child Pauses	13.50	11.50	8.50	4.75	0.15
Parent Switching Pauses	24.00	20.75	14.00	5.25	0.62
Child Switching Pauses	22.00	20.00	12.00	5.25	0.50
Vocal Durations[c]					
Parent Loudness	0.25	0.29	0.12	0.06	0.19
Child Loudness	0.09	0.16	0.05	0.06	0.73
Parent Switching Pause	1.13	1.82	1.91	0.87	0.21
Motor Durations[c]					
Near	43.23	54.48	53.21	59.88	0.38
Parallel/Oblique	11.19	32.47	25.93	27.41	0.52

Note. Medians and interquartile ranges (IQR) of each behavioral variable by group are shown. [a]Wilcoxon rank-sum test. [b]Frequencies denote the number of occurrences. [c]Durations denote the amount of time spent in seconds.

Appendix

The Vocal-Motor Coding System (VMCS)

The Vocal-Motor Coding System (VMCS) was developed to identify the discrete motoric and vocal elements of a social interaction between a parent and a child. The VMCS is comprised of two coding schemes—the motoric behavioral coding scheme (see Table A1) and the vocal behavioral coding scheme (see Table A2). Motoric behaviors are nominally rated across dimensions of *Proximity*, *Orientation*, and *Object-Touch*, while vocal behaviors are rated across dimensions of *Loudness* and *Rhythmicity*. We recommend using a video-coding tool, such as Datavyu (Datavyu Team, 2014), to code each onset of a motoric behavior, and an audio-editing program, such as Audacity® (Audacity Team, 2017), to clean and manipulate audio data for vocal behavioral coding.

Motoric Behavioral Coding Scheme

The motor component of the VMCS was developed to capture three physical dimensions of dyadic social interaction. Each dimension is further outlined below.

Proximity. *Proximity* is rated as either "Near" or "Far," depending on whether the child is within an adult arm's reach of the parent. In most dyadic play interactions, the parent does not move from his or her space on the floor, and the child moves in and out of arm's reach of the parent. The average length of a human adult arm ranges from 27 to 30 inches. As it is not possible to determine the exact length of space between parent and child in most video coding tools, we recommend that the rater uses his or her best judgment to determine whether the parent and child are within arm's reach. If the parent and child are at or within the parent's reach of each other, code the Proximity of the dyad as "Near." If the parent requires minimal effort to reach for his or her child (e.g., rising a couple inches off the ground in a cross-legged or kneeling position to reach the child or leaning forward slightly from his or her seated position to reach the child), the parent and child would be coded as "Near" each other. The Proximity of the dyad is rated as "Far" if the parent and child are visibly out of an adult arm's reach of each other. Each onset of "Near" and "Far" is coded regardless of the behavior's duration.

Orientation. *Orientation* is rated as "Toward," "Parallel/Oblique," or "Away," depending on the alignments and positioning of the parent's and child's chest regions respective of each other. The Orientation of a dyad is rated as "Toward," if the parent's and child's chest regions are directly facing each other during the interaction, regardless of the distance between them. The Orientation is rated as "Parallel/Oblique" if the parent and child are interacting with each other side-by-side or while positioned within 90 degrees of each other. Rate the Orientation of a dyad as "Away" if the chest regions of the parent and child are not facing each other because of their positions in space with respect to each other. Refer to Figure A1 for example illustrations of each Orientation code.

Object-Touch. *Object-Touch* is rated each for the parent and child in an interaction and signifies whether the parent or child is touching the shared object. The onset of each instance of Object-Touch is coded with "Yes," and the offset with "No," for each social partner.

Vocal Behavioral Coding Scheme

The vocal component of the VMCS captures two vocal dimensions of dyadic social interaction, which are further detailed below. Audio data of each dyad is split in Audacity® into two tracks, one designated each for the parent and the child's vocalizations. Each track is then removed of extraneous noise and then silenced in segments where no vocalizations are heard. The final, cleaned version of the dyad's audio data is then passed through the Parent-Child Vocalization Analyzer, an automated program written in MATLAB (R2017a; The Mathworks,

2017) that quantifies each instance of a vocal state into codes of “0” for silence, “1” for the parent’s vocalization, and “2” for the child’s vocalization.

Loudness. The rating of the parent or child’s *Loudness* is either “Loud” or “Not Loud” and also performed in the Parent-Child Vocalization Analyzer. A rating of “Loud” or “Not Loud” depends on whether the amplitude of a vocalization signal is above or below a specific threshold that is calculated each for the parent and the child. This partner-specific threshold is the value obtained from dividing the *range* of amplitudes of the partner’s vocalizations by two. Each sample of a vocalization in the audio data whose amplitude is greater than this value is coded as “Loud.” Conversely, each sample of a vocalization whose amplitude is smaller than this value is coded as “Not Loud.” For example, for a child’s vocalizations whose total range of waveform amplitudes is equivalent to 0.94, his or her *Loudness* threshold would be 0.47. Refer to Figure A2 for a graphical representation of a partner-specific threshold.

Rhythmicity. Vocal rhythm is comprised of *vocalizations*, *pauses*, and *switching pauses* of the parent and child each. This coding scheme is adapted from the vocal rhythm variables described by Feldstein et al. (1982):

> A *vocalization* is a segment of sound (speech) uninterrupted by any silence that can be heard by the unaided human ear and that is uttered by the speaker who has the speaking turn. A *pause* is an interval of joint silence bounded by the vocalizations of the speaker who has the turn, and is thereby credited to him or her. A *switching pause* is an interval of joint silence that is initiated by the speaker who has the turn and terminated by the other speaker, who thereby obtains the turn. (Feldstein et al., 1982, p. 454).

Refer to Figure A3 for a diagrammatic representation of a parent-child interaction sequence.

Table A1

Motoric Behavioral Codes of the Vocal-Motor Coding System

Behavioral Dimension	Dyadic Code	Parent Code	Child Code
Proximity	Near (N) Far (N)		
Orientation	Toward (T) Parallel/Oblique (P) Away (A)		
Object-Touch		Yes (Y) No (N)	Yes (Y) No (N)

Table A2

Vocal Behavioral Codes of the Vocal-Motor Coding System

	Parent	Child
Loudness	Loud Not Loud	Loud Not Loud
Rhythmicity	Vocalization Pause Switching Pause	Vocalization Pause Switching Pause

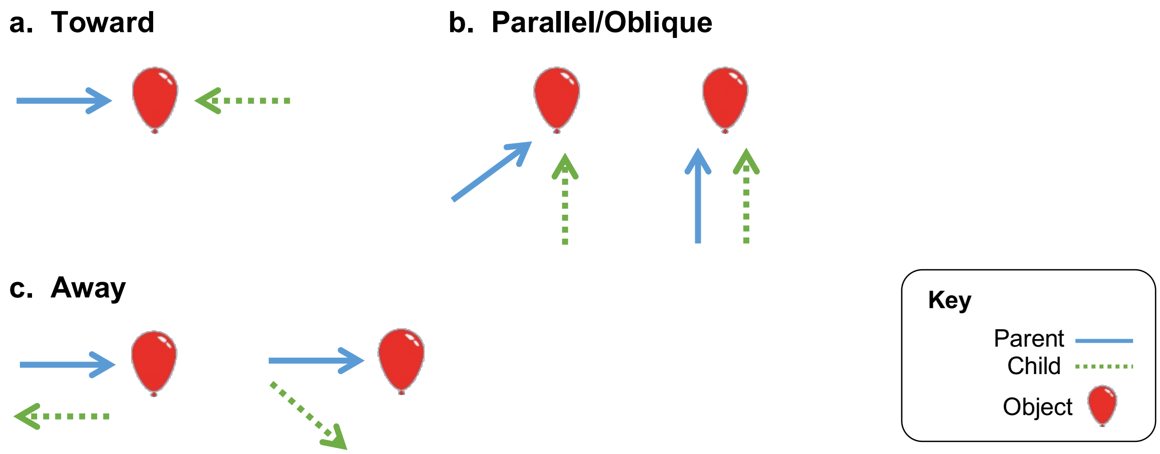

Figure A1. Example scenarios of dyadic Orientation codes: (a) *Toward*; (b) *Parallel/Oblique*; and (c) *Away*. (a, b) In each scenario of *Toward* and *Parallel/Oblique*, the directions in which the parent and child are facing converge towards each other or the shared object. (c) In each scenario of *Away*, the direction of one partner does not converge with that of the other; rather, they are positioned back-to-back, or with one partner's front facing the other's back.

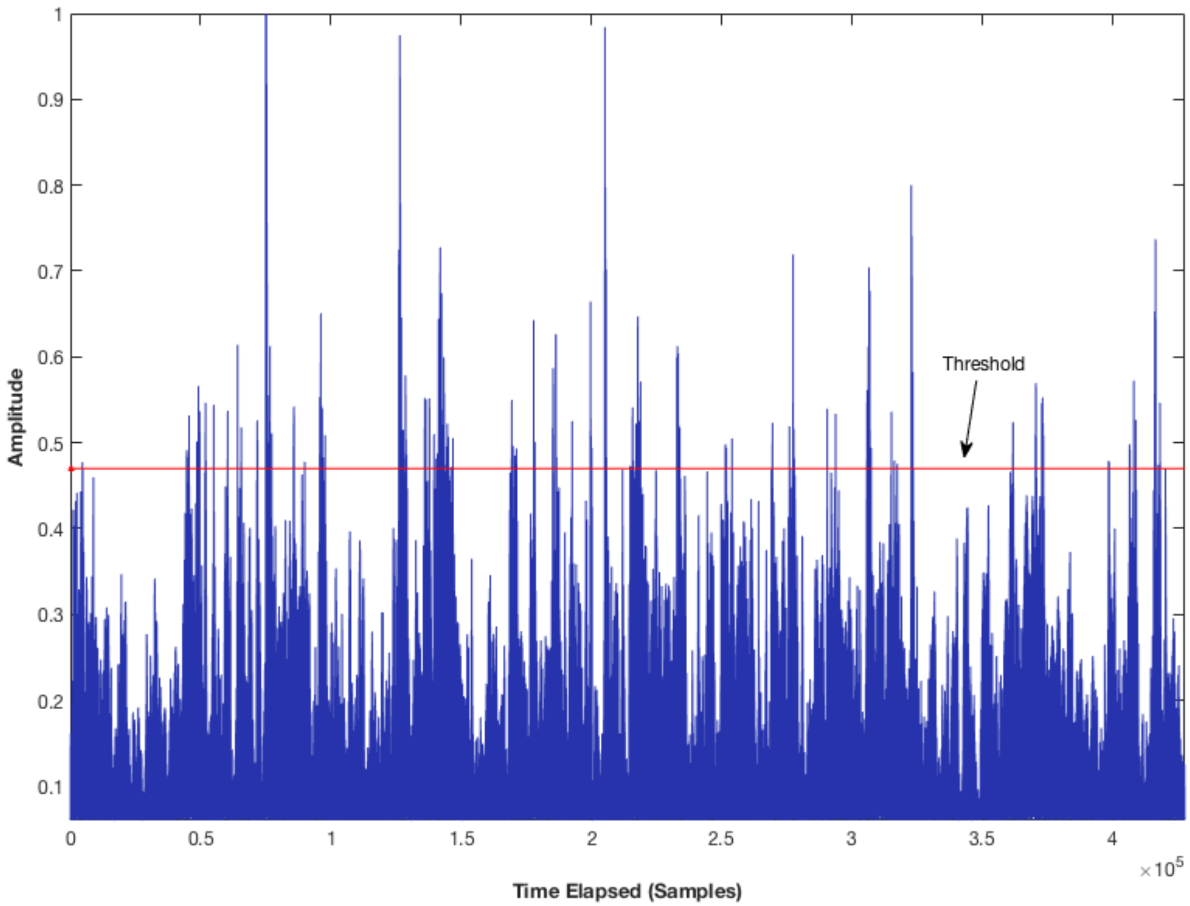

Figure A2. Graphical example of vocalizations sampled from a child's audio track (n = 428,353, f_s = 32 kHz) and the "loudness threshold" (y = 0.47). All vocalization signals with an amplitude ≥ 0.47 are coded as "Loud," and all vocalizations with an amplitude < 0.47 are coded as "Not Loud."

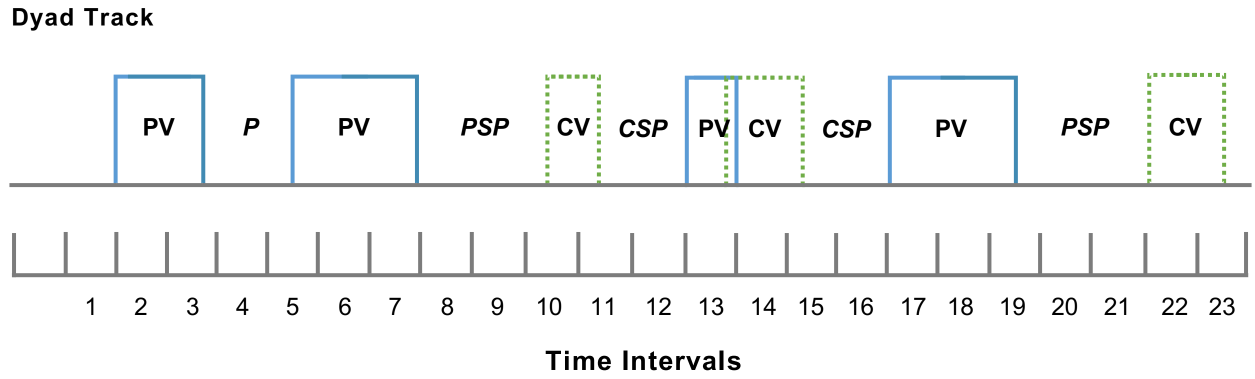

Figure A3. Diagrammatic representation of a dyadic interaction sequence. *PV* = parent vocalization; *CV* = child vocalization; *P* = pause; *PSP* = parent switching pause; and *CSP* = child switching pause. The time axis represents intervals of vocalizations at a track-specific sampling rate (e.g., 32,000 samples/sec)