A New Measure of Child Vocal Reciprocity in Children with Autism Spectrum Disorder

Amy L. Harbison, Tiffany G. Woynaroski, Jon Tapp, Joshua W. Wade, Anne S. Warlaumont, and Paul J. Yoder

Children's vocal development occurs in the context of reciprocal exchanges with a communication partner who models "speechlike" productions. We propose a new measure of child vocal reciprocity, which we define as the degree to which an adult vocal response increases the probability of an immediately following child vocal response. Vocal reciprocity is likely to be associated with the speechlikeness of vocal communication in young children with autism spectrum disorder (ASD). Two studies were conducted to test the utility of the new measure. The first used simulated vocal samples with randomly sequenced child and adult vocalizations to test the accuracy of the proposed index of child vocal reciprocity. The second was an empirical study of 21 children with ASD who were preverbal or in the early stages of language development. Daylong vocal samples collected in the natural environment were computer analyzed to derive the proposed index of child vocal reciprocity, which was highly stable when derived from two daylong vocal samples and was associated with speechlikeness of vocal communication. This association was significant even when controlling for chance probability of child vocalizations to adult vocal responses, probability of adult vocalizations, or probability of child vocalizations. A valid measure of children's vocal reciprocity might eventually improve our ability to predict which children are on track to develop useful speech and/or are most likely to respond to language intervention. A link to a free, publicly-available software program to derive the new measure of child vocal reciprocity is provided. ***Autism Res** 2018, 11: 903–915*. © 2018 International Society for Autism Research, Wiley Periodicals, Inc.

Lay Summary Children and adults often engage in back-and-forth vocal exchanges. The extent to which they do so is believed to support children's early speech and language development. Two studies tested a new measure of child vocal reciprocity using computer-generated and real-life vocal samples of young children with autism collected in natural settings. The results provide initial evidence of accuracy, test-retest reliability, and validity of the new measure of child vocal reciprocity. A sound measure of children's vocal reciprocity might improve our ability to predict which children are on track to develop useful speech and/or are most likely to respond to language intervention. A free, publicly-available software program and manuals are provided.

Keywords: reciprocity; vocalizations; automated vocal analysis; LENA; preschool; preverbal; autism

Introduction

Typically developing children naturally engage in reciprocal vocal exchanges with an adult communication partner who serves as a model of more speechlike vocalizations [Gros-Louis, West, & King, 2014; Kuhl, 2003]. For example, a child may initiate an exchange by producing a vocalization to which an adult vocalizes or verbalizes. The child may then produce another vocalization in response to the adult's vocal response. That is, the concept requires that the child's vocalization does not just occur after the prior adult vocal response, but rather that it occurs because the child attends to and is affected by the prior adult vocalization. We call such sequences *vocal reciprocity*. Such back and forth vocal responding has been theorized to create a social feedback loop that promotes speech and language development [Warlaumont, Richards, Gilkerson, & Oller, 2014]. Specifically, contingent adult vocal responses may increase the child's attention to and eventual emulation of the speechlike qualities of the adult's vocalizations [Goldstein, et al., 2010]. Emulation, in this

From the Department of Special Education, Vanderbilt University, Nashville, TN (A.L.H.); Department of Hearing and Speech Sciences, Vanderbilt University Medical Center, Vanderbilt Kennedy Center, Vanderbilt Brain Institute, Nashville, TN (T.G.W.); Vanderbilt Kennedy Center, Vanderbilt University Medical Center, Nashville, TN (J.T.); Department of Mechanical Engineering, Vanderbilt University, Nashville, TN (J.W.W.); Department of Communication, University of California, Los Angeles, CA (A.S.W.); Department of Special Education, Vanderbilt University, Nashville, TN (P.J.Y.)

Received June 13, 2017; accepted for publication February 19, 2018

Address for correspondence and reprints: Paul Yoder, Department of Special Education, Vanderbilt University, 110 Magnolia Circle, Office 416B, Nashville, TN, 37203. E-mail: paul.yoder@vanderbilt.edu

Published online 6 March 2018 in Wiley Online Library (wileyonlinelibrary.com)
DOI: 10.1002/aur.1942

context, does not necessarily mean the child immediately or exactly imitates adult vocal models. Instead, emulation, as used there, means the child increasingly uses vocalizations with speechlike characteristics.

Speechlikeness of Vocalizations in Children with Autism Spectrum Disorder (ASD)

As a group, children with autism spectrum disorder (ASD) tend to produce fewer speechlike vocalizations relative to their peers [Paul, Chawarska, Cicchetti, & Volkmar, 2008]. Children with ASD, however, are highly heterogeneous in the speechlikeness of their vocalizations, and individual differences in the speechlikeness of vocalizations appear to serve as a useful index of progress in native language acquisition for children with ASD in the early stages of language development [Woynaroski et al., 2017; Yoder, Watson, & Lambert, 2015]. Identifying factors that correlate with or predict speechlikeness of vocalizations may help us to better understand the heterogeneity that we see in speechlikeness across children with ASD. Such predictors of speechlikeness of vocalizations may also improve the accuracy of our prognostic statements regarding speech and language development in children with ASD. Finally, identifying factors that explain variability in speechlikeness of vocalizations may facilitate personalized treatment planning for those children whose vocal development does not appear to be on track.

Vocal Reciprocity in Children with ASD

Children with ASD are likely to experience fewer reciprocal vocal exchanges than developmental- or chronological-age matched peers for several reasons. For example, children with ASD on average produce a smaller number of vocalizations to which parents may respond than do developmentally matched peers [Patten et al., 2014]. Relative to typically developing peers, children with ASD also produce a smaller proportion of vocalizations that are speechlike, the type of child vocalization that is most likely to elicit language-facilitating responses from adult communication partners [Gros-Louis et al., 2014; Paul et al., 2008]. When adults do respond to child vocalizations, children with ASD may show less attention to, and less preference for, child-directed speech than developmentally matched peers [a "tuning-in" deficit; Baranek, 1999; Klin, 1991; Lord, 1995]. Even when they do attend to the model, children with ASD may have a deficit in emulating the characteristics of adults' speech [a "tuning-up" deficit; Shriberg, Paul, Black, & van Santen, 2011].

Individual differences in speechlikeness of vocalizations of young children with ASD may be explained by variation in children's generalized tendency to engage in reciprocal vocal exchanges. It is proposed here and other places [Warlaumont et al., 2014] that child vocal reciprocity may reflect the extent to which children with ASD attend to and process adult speech. According to the social feedback theory, child vocal reciprocity should be associated with the speechlikeness of children's vocal communication during the preverbal and early stages of language development.

Daylong Vocal Samples as a Measure of Child Vocal Reciprocity

Measurement of vocal reciprocity in children with ASD poses a challenge. Estimates of child vocal reciprocity from short vocal samples are unlikely to have high test-retest reliability. Measures with poor test-retest reliability are limited in their potential utility for explaining individual differences (variance) in outcomes of interest, such as speechlikeness [Crocker & Algina, 1986]. Stable measures of child vocal reciprocity would have the potential to correlate more strongly with speechlikeness of vocalizations [Yoder & Symons, 2010]. Although not specifically focused on vocal exchanges, Staubitz and Lloyd [2016] found that only long (i.e., exceeding 300 min) observations of unstructured interactions produced acceptably high test-retest reliability for estimates of complex interactional sequences.

Daylong samples of child and adult vocalizations could provide the amount of data needed to provide stable estimates of child vocal reciprocity. Collecting such extensive data is potentially feasible because existing Language ENvironment Analysis (LENA) technology affords easy collection of daylong samples in naturalistic settings [LENA Research Foundation, 2016]. The LENA system can be used to collect and automatically segment lengthy audio streams and to classify acoustic events, including target child vocalizations (CV) and nearby adult vocalizations (AV). Thus, daylong vocal samples serve as a good candidate for measurement of vocal reciprocity in young children with ASD.

Important Considerations regarding Putative Metrics of Child Vocal Reciprocity

There are a number of points to consider in selecting a candidate metric of child vocal reciprocity to be derived from daylong vocal samples. By *metric*, we mean the type of score derived from a measure. A content-valid metric of child vocal reciprocity must tap the bidirectional influence between the child and their adult communication partner/s. Thus, a three-event sequence between child and adult vocalizations is central to our concept of child vocal reciprocity. When applied to a child's vocal reciprocity, this three-event sequence is best conceptualized as a child vocalization followed by an adult vocalization followed by a subsequent child vocalization (CV→AV→CV).

As shown in past studies, though, the *number* of times one might observe a CV→AV→CV sequence is highly likely to be influenced by the mere frequency of child and/or adult productions [Yoder & Symons, 2010]. That is, as a child and/or adult vocalize more frequently, we would be more likely to observe the CV→AV→CV sequence by chance alone. An example of a chance occurrence of the second CV in a CV→AV→CV sequence occurs when the child is not responding to the preceding adult vocal response, yet the CV→AV→CV sequence occurs. Such an event can occur when an adult inserts his or her vocal turn between two child preverbal vocalizations that are not directed to the adult. If the association between a putative measure of child vocal reciprocity and speechlikeness of vocalizations becomes nonsignificant when controlling for the frequency of CV, the simplest interpretation of this putative measure of child vocal reciprocity is that it mainly reflects child vocalizations, and does not capture a dyadic process. What is needed is a metric that captures the three-event CV→AV→CV exchange between a child and adult/s in a manner that controls for chance-level sequencing of the three events.

The Potential to Improve upon Previously Developed Metrics of Child Vocal Reciprocity

A number of investigations have used daylong vocal samples to compute Conversational Turn Count/s (CTC), which index the frequency of two-event sequences of child and adult vocalizations that occur in close temporal proximity [Ganek & Eriks-Brophy, 2017; Dunst, Hamby, Trivette, Prior, & Derryberry, 2013]. CTC indices (total, adult-initiated, and/or child-initiated CTCs) can be derived from daylong vocal samples using commercially available LENA software. However, CTC indices are unlikely to reflect the bidirectional nature of child vocal reciprocity because (a) they are only two-event sequences, and (b) they do not control for, and thus are very likely influenced by, the chance sequencing of child and adult vocalizations. Another recent study that used daylong vocal samples focused on three-event sequences to assess the relationship between adult vocal responses and subsequent child vocalizations in children with ASD [Warlaumont et al., 2014], focusing on different types of child vocalizations that followed an adult vocal response (cry/laugh/vegetative vs. speech-related). This study represents an advance in tapping the bidirectional nature of the exchange between a child and adult(s) because it is based on CV→AV→CV sequences. However, a thorough examination of the index with consideration of potential confounding variables, such as the overall frequency of CV or AV, was not conducted. Moreover, that measure has not been validated in relation to other measures of child communication ability. Thus, prior to the present study, we still did not have a measure of vocal reciprocity that (a) captured the three-event CV→AV→CV sequence and (b) controlled for the chance sequencing of CV and AV events.

The Reciprocal Vocal Contingency Score as a Novel Metric of Child Vocal Reciprocity

In the present study, we put forth a metric that may improve upon past attempts at quantifying child vocal reciprocity—the *Reciprocal Vocal Contingency* (i.e., *RVC*) score. The RVC score is a sequential association between a child's vocal response to an immediately preceding adult vocal response. It captures the three-event CV→AV→CV sequence that taps the bidirectional exchange between a child and adult/s. Even more importantly, the RVC was designed to account for chance probability of sequencing. A *positive* sequential association, applied to our example, means that child vocalizations occur after adult vocal responses to child vocalizations more than expected by chance sequencing of child and adult vocalizations [Gottman & Roy, 1990]. The higher the sequential association (i.e., the closer to 1), the more the CV→AV→CV sequence exceeds chance estimates. We call this index RVC because another term used for a sequential association is a *contingency* [Lloyd, Kennedy, & Yoder, 2013]. Although the causal influence of CV→AV on CV is not certain in positive sequential associations, the positive sequential association of CV→AV→CV provides correlational evidence that the preceding adult vocal response influences the child's vocal response while ruling out a primary alternative explanation to the association (i.e., that the sequence occurred by chance due to the mere frequency of child and/or adult vocalizations).

Aims

Our primary aim was to examine the psychometrics of the newly-proposed RVC score. We did so in two studies. In the first study, we tested the accuracy of this novel metric in the context of a simulation that utilized computer-generated event sequences, which can be thought of as simulated vocal samples. In the second study, we examined the test-retest reliability and construct validity of RVC using real-life vocal samples of young children with ASD.

A scientifically useful measure of children's vocal reciprocity has evidence of high reliability and construct validity. Classical measurement theory tells us that the validity of RVC as a correlate of speechlikeness of vocal communication in children with ASD is limited by RVC's test-retest reliability [Crocker & Algina, 1986]. At present, we cannot use criterion-related validity to assess the construct validity of RVC because there is no gold standard measure of child vocal reciprocity. A gold standard measure is one that has been extensively validated and widely accepted measure of the same construct. It is possible in these instances, however, to test whether a new measure or metric shows lawful associations with measures of other, logically related constructs, but not with measures of constructs for which the new measure should not be related [Cronbach & Meehl, 1955; Yoder & Symons, 2010]. We specifically evaluate whether RVC shows a theoretically-predicted association with speechlikeness of vocal communication in our sample of children with ASD. We also sought to show RVC is unrelated to variables that theory suggests it should be unrelated (i.e., chronological age, level of cognitive impairment, and parent's formal education level).

Study 1

Introduction

In Study 1, we assessed the accuracy of our metric of vocal reciprocity, RVC. The only way to judge the absolute accuracy of RVC was to test whether it produced a score that was nearly equal to a known sequential association for the CV→AV→CV sequence. However, the only currently known way to ensure the magnitude of the sequential association in a vocal sample was a particular value was to generate a large number of simulated vocal samples with randomly sequenced CV and AV events. For each simulated vocal sample, we obtained an RVC score. Even when computed on randomly sequenced vocal samples, we would expect some of the RVC scores from *individual* vocal samples to depart from zero to some degree. The mean RVC from very large number of simulated vocal samples with randomly sequenced CV and AV events should be at or very near zero, though, if the RVC is an accurate index of sequential association of CV→AV→CV.

In the simulated vocal samples with randomly sequenced and timed CV and AV events, we could also test whether the method of estimating the frequency of chance occurrence of the CV→AV→CV sequence was accurate (i.e., nearly perfectly correlated with the recorded number of CV→AV→CV in the set of randomly sequenced events). Demonstrating the accuracy of the RVC score, and of the metric we proposed to use as an index of chance sequencing of CV→AV→CV, improves the basis for interpreting the RVC index and provides a foundation for validity testing in an empirical sample of children with ASD (i.e., Study 2).

Methods

Sequential analysis method. We used a sequential analysis approach called *event lag with contiguous pauses* to analyze the simulated vocal samples of child and adult vocal productions [Lloyd, Yoder, Tapp, & Staubitz, 2015]. This method involves identifying the events of interest (i.e., CV and AV events) in each sample and stripping out all other event types. Additionally, when events of interest do not occur for a given length of time, fixed-duration pauses are inserted between events of interest and analyzed as if they were events. In this study, contiguous 2 sec pauses were inserted when neither CV nor AV occurred for at least 2 sec. We utilized 2 sec pauses because this is a commonly documented pause duration in interactions between adults and young children in the early stages of language development [e.g., Brundin, Rödholm, & Larsson, 1988; Gros-Louis, West, Goldstein, & King, 2006; Northrup & Iverson, 2015].

The event lag with contiguous pauses method was selected because decisions regarding which non-key events (i.e., those other than the events in the key sequence) should be analyzed are usually underjustified but can greatly influence the strength and the direction of the sequential association [Lloyd et al., 2015; Yoder & Symons, 2010]. An extant simulation study has demonstrated that the event lag with contiguous pauses method resolves the problem of having to justify which non-key events should be included in the computation of a sequential association [Lloyd et al., 2015]. The event lag with contiguous pauses method has also been shown to be more accurate and less correlated with the frequency of chance occurrence of the sequence of interest than three other sequential analysis methods [i.e., time window, concurrent interval, and event lag without pauses; Lloyd et al., 2015]. Further detail about the event lag with contiguous pauses method is found in Lloyd et al. [2015].

Data generation method. We sought to generate 5,000 1-hr simulated vocal samples with a mean sequential association of zero. To model the effect of stripping out non-key events (as routinely done in the event lag with contiguous pauses method of sequential analysis), we generated simulated vocal samples with only two key events: CV and AV. To generate random frequencies for each event type, we first randomly selected a number from a uniform distribution with an empirical minimum of 10 and an empirical maximum of 2500 for each key event. We judged that this range would produce sufficient frequency of occurrence of key events to quantify a nonzero contingency if it were present in the data, while allowing sufficient basis for generalization across a wide range of probabilities of the key event types. For example, if 100 CV and 500 AV were selected for a simulated vocal sample, the computer generated a total of 600 events. Next, we randomized the sequence of key events by randomly assigning

		Event 3: CV	Event 3: (not CV)	Marginals: Events 1 and 2
Events 1 and 2	[CV → AV]	[CV → AV] → CV *A**	[CV → AV] → (not CV) *B*	Total [CV → AV] *A + B*
	(not [CV → AV])	(not [CV → AV]) → CV *C*	(not [CV → AV]) → (not CV) *D*	Total (not [CV → AV]) *C + D*
	Marginals: Event 3	Total (Event 3 = CV) *A + C*	Total (Event 3 = not CV) *B + D*	Total # of events *A + B + C + D*

Figure 1. 2 × 2 contingency table used to tally three-event sequences of child and adult vocalizations. CV = child vocalization. AV = adult vocalization. The → symbol = followed by. Cells comprising the 2 × 2 table are highlighted. Cell labels (A, B, C, D) are italicized in the bottom center of each cell. Cell A represents the three-event sequence of special interest, wherein an initial child vocalization is followed by an adult production that is followed by a subsequent child vocalization, and is indicated here with an asterisk. Cell B represents a two-event sequence, wherein an initial child vocalization is followed by an adult production, but this is NOT followed by a subsequent child vocalization. Cell C represents a child vocalization that occurs in the absence of a preceding two-event sequence, wherein an initial child vocalization is followed by an adult production. Cell D represents instances wherein there is neither an initial child vocalization followed by an adult production nor a subsequent child vocalization.

without replacement each event to one of 3,600 (i.e., the number of seconds in 1 hr) positions within the simulated vocal sample, which is analogous to time of occurrence. Finally, we inserted 2-sec pauses when neither child nor adult vocalizations had occurred for at least 2 sec to enable quantification of immediacy (i.e., the extent to which one key event, such as CV, immediately follows another key event or sequence of key events, such as CV→AV), while maintaining equal weighting of events and pauses in the analysis. Multiple contiguous pauses were inserted when the key events were absent for at least twice the duration of the fixed pause (i.e., for at least 4 sec).[1]

Data reduction method. The resulting simulated vocal samples were automatically analyzed by a computer program to tally three-event sequences (i.e., each sequence of three consecutive events) into one of four cells described by a 2 × 2 contingency table with row and column labels as indicated in Figure 1.

The reader is asked to carefully note the row and column labels. The antecedent unit (i.e., the first row) in this 2 × 2 table is defined by the CV→AV sequence. In Table 1, we present a truncated timed-event sequence that includes CV and AV events with contiguous pauses with an indication of how three-event sequences are tallied into the 2x2 table cells.

[1]When the pause duration is not evenly divided by 2 (an example of this occurs at seconds 9 through 11 in Table 1), the number of pause events inserted will not precisely represent the pause duration. The event lag with contiguous pauses approach results in analyzing the session as an event stream in which the number of events does not exactly represent the duration of the observation. There is no evidence, however, that analyzing pauses as events with equivalent weight as CV and AV biases the results in any way.

Special attention to the definition of the A cell is warranted. The A cell count is the observed frequency of CV→AV→CV. As the events were randomly sequenced in this simulation study, this value should be highly similar to the computed chance frequency of CV→AV→CV.

For further explanation of the logic behind interpreting 2 × 2 tables in this manner and for a rationale of the overlapping window approach to tallying event sequences into 2 × 2 tables, see Bakeman and Gottman [1997]. Cell values for the 2 × 2 table were computed for each event stream and used to compute the following variables.

The probability of chance occurrence of CV→AV→CV was computed via formula (1).

$$P_{\mathrm{cv}} \times P_{\mathrm{cv}\rightarrow\mathrm{av}} \tag{1}$$

Using the cell values labeled A through D in the 2 × 2 table, P_{cv} is the probability of child vocalization ([A + C]/[A + B+C + D]) and $P_{\mathrm{cv}\rightarrow\mathrm{av}}$ is the probability of CV→AV sequences ([A + B]/[A + B+C + D]).

The RVC was computed using the formula (2):

$$P_{\mathrm{cv},\,(\mathrm{cv}\rightarrow\mathrm{av})} - P_{\mathrm{cv},\,\sim(\mathrm{cv}\rightarrow\mathrm{av})} \tag{2}$$

Where $P_{\mathrm{cv},\,(\mathrm{cv}\rightarrow\mathrm{av})}$ is the probability of CV given the prior occurrence of the two-event antecedent unit containing CV→AV and $P_{\mathrm{cv},\,\sim(\mathrm{cv}\rightarrow\mathrm{av})}$ is the probability of CV given the prior occurrence of any other type of two-event antecedent unit. Positive values of RVC support an inference that adult vocal responses to child vocalizations elicit the immediately following child vocalizations.

The metric we used to quantify the sequential association of CV→AV→CV, which has also been called an operant contingency value [Hammond, 1980; Martens, Gertz, Werder, Rymanowski, & Shankar, 2014], is

Table 1. Tallying Method for a Truncated Timed-Event Sequence of CV and AV Onsets with Contiguous Pauses Inserted

Seconds	Events	Overlapping 3-event sequence	2 × 2 cell label
1	CV		
2	AV		
3	AV	1	B
4	CV	2	C
5	AV	3	D
6	CV	4	A*
7–8	P	5	D
9–11	P	6	D
12	CV	7	C
13–14	P	8	D
15	AV	9	D

Note. CV = child vocalization onset, AV = adult vocalization onset, P = pause. *A denotes the three-event sequence of interest.

derived from an exhaustive and mutually exclusive accounting of the observed sequence of behaviors and has been shown previously to be independent from chance sequencing of events [Lloyd et al., 2013]. When contingency tables are sparse, operant contingency values are a more accurate estimator of contingency than other putative indices of contingency [Lloyd et al., 2013]. Operant contingency values are also more conceptually related to operant contingency theory than other indices of sequential association. More information on the mathematics underlying quantification of sequential associations is found in Martens et al. [2014].

Results of Study 1

Is RVC an Accurate Quantification of Child Vocal Reciprocity?. The mean RVC from the simulated vocal samples was .0002 ($SD = .058$), 95% CI [−.0014, .0018]. Cohen's d for the difference from 0 was .003, $t(4999) = .25$, $P = .8$. As the simulated vocal samples were generated using a method that produced a random frequency, sequencing, and timing of events, we can be assured that the mean contingency of the underlying distribution was zero. Because the empirical mean of the RVC was almost exactly the known mean of the underlying distribution, we have evidence that the RVC score accurately reflects the contingency of the CV→AV→CV sequence, which we are proposing as an index of child vocal reciprocity.

To test the accuracy of the method used to estimate the frequency of chance sequencing, we computed the correlation of the estimated chance probability of the key sequence (i.e., derived from formula [1]) with the probability of the observed CV→AV→CV sequence (i.e., the A cell count/total events) from the simulated vocal samples. The Pearson product moment correlation was .996.

Discussion of Study 1

Study 1 demonstrated that the RVC is an accurate estimate of sequential association for the CV→AV→CV sequence, meaning that this novel metric can be used to quantify the sequential association of this three-event sequence of interest, which controls for chance probability of sequencing. Our estimate of the probability of chance sequencing of CV→AV→CV was also accurate. As with all simulations, the results are only applicable to the real world if the simulation data are generated using realistic constraints. At present, absolute accuracy, and thus possible bias, of RVC can only be judged relative to a known mean contingency, and a known mean contingency can only be known when it is zero (i.e., generated from a population of simulated vocal samples with randomly sequenced and timed events). In a set of real vocal samples from a sample of children with ASD, the mean contingency is likely to be nonzero. Thus, Study 2 draws upon real data from children with ASD to further test the psychometrics of the newly developed index of child vocal reciprocity.

Study 2

Introduction

To begin the process of determining whether RVC is psychometrically sound when applied to real vocal samples of children with ASD, we analyzed daylong vocal samples from children with ASD who were in the early stages of expressive language development. We used computer software to compute RVC scores from the acoustic events in the audio recording of the daylong audio-recorded samples to demonstrate the feasibility of using the RVC in low-cost research and ultimately, pending sufficient validation of the metric, in clinical practice. We then assessed the short-term test-retest reliability of RVC and tested the construct validity of RVC in the sample of children with ASD. The social feedback loop theory supported our prediction that child vocal reciprocity would be correlated with speechlikeness of vocal communication in this sample.

We focus on consonant inventory as the measure of speechlikeness of vocal communication in this study because consonant inventory measured during the preverbal period was previously identified as a predictor of expressive language growth in children with ASD even after controlling for eight other theoretically and empirically justified putative predictors of expressive language [Yoder et al., 2015]. Consonant inventory also correlates highly with two other measures of speechlikeness of vocalizations in children with ASD [Woynaroski et al., 2017].

Table 2. Participant Characteristics in Study 2

Characteristic	*M*	*SD*	10th%	90th%	Median
Chronological age (months)	41	5	28	47	41
MSEL developmental quotient	25	10	14	38	25
ADOS diagnostic algorithm score	24	4	17	28	24
MB-CDI expressive raw score	50	68	0	191	17
MB-CDI receptive raw score	85	106	0	265	24

Note. MSEL = Mullen Scales of Early Learning [Mullen, 1995]. ADOS = Autism Diagnostic Observation Schedule [Lord et al., 2000]. MB-CDI = MacArthur-Bates Communicative Development Inventory: Words and Gestures form [Fenson et al., 2007].

Methods

Participants. Participants in the present study were 21 preschoolers with ASD (18 male, 3 female; chronological ages 29 to 47 months old) from a larger correlational study [Yoder et al., 2015]. The subset of participants included in the present study were the children in the larger study for whom (a) at least one daylong audio recording of the child's language environment and (b) a concurrent Communication and Symbolic Behavior Scales – Developmental Profile Behavior Sample [CSBS-DP; Wetherby & Prizant, 2002] were collected.

Diagnoses of ASD were based on the Autism Diagnostic Observation Schedule [Lord et al., 2000] and the clinical judgment of an experienced diagnostician using the Diagnostic and Statistical Manual of Mental Disorders criteria [4th ed., text rev.; American Psychiatric Association, 2000]. The Mullen Scales of Early Learning [Mullen, 1995] was also administered at entry to the larger study to further characterize the sample. Participants were in the early stages of language development, but their language skills at the time of the recordings were somewhat variable. Sixteen of the children (i.e., 76%) were considered preverbal, based on their parents' report that the children produced no more than 20 words on the MacArthur-Bates Communicative Development Inventory: Words and Gestures form [MB-CDI; Fenson et al., 2007]. Further information about the participants' language and other characteristics is summarized in Table 2.

The parent who served as the primary caregiver for each participant self-reported the highest level of formal education that they had achieved by checking one of nine educational levels on a demographic questionnaire. Parents reported achieving a mean formal education level of 1–2 years of college or technical school.

Measurement of child vocal reciprocity. Daylong vocal samples were collected in participants' natural environments using a small digital recording device that is part of the LENA system. More detail about the LENA system can be found on the website for the LENA Research Foundation [2016]. Parents were instructed to turn on the recorder when their child woke up in the morning, to place the recorder on their child in the chest pocket of a specially designed vest provided by the research team, and to allow the recorder to run continuously for a full day. There were no constraints placed on the day(s) the digital recorder was worn and turned on. All parents involved in the study were able to comply with the instructions regarding data collection.

Audio recordings were collected on two consecutive days for 86% ($n = 18$) of the participants. Participants with two vocal samples were included in the estimates of short-term test-retest reliability of RVC. For the test of validity, RVC scores were averaged across the 2 days' samples, if available, and were derived from one daylong sample if not (14%, $n = 3$). Recordings averaged 14.9 hr ($SD = 2.5$ hr) each. The average number of hours recorded per participant was 28.3 hr ($SD = 7.1$ hr). When recorders were returned, research staff transferred audio files from the recorders directly to a computer for analysis.

We used the LENA Pro software to segment acoustic events and classify them into multiple categories including key child (i.e., child wearing the LENA recorder) and near adult productions. Physical proximity of "near adult productions" is classified automatically by LENA algorithms according to the intensity or "loudness" of the adult-labeled vocalization compared to a minimum average loudness observed when speakers are within about 6 ft of the microphone, which in this case was worn by the key child. "Far" adult vocalizations were excluded. Segments labeled as likely being produced by the key child were then further classified into speech related utterances (i.e., speech related vocalizations of at least 50 ms duration bounded by sounds of other source types or silence for more than 300 ms) versus fixed signals (cries) or vegetative sounds (such as burps). The timing and classification of each acoustic event is available via an output file called an Interpreted Time Segments (ITS) file that can be obtained from the LENA Pro software. More information about the reliability of classification of audio segments relative to human coding, as well as the content of ITS files, is available on the LENA Research Foundation's website [Xu, Yapanel, Gray, & Baer, 2008].

The ITS files from the LENA Pro software were then used as input for a custom-made software program [i.e., Contingencies from LENA data; Yoder, Wade, Tapp, Warlaumont, and Harbison, 2016], that performed two more steps necessary for computing RVC. First, all events except for key child speech related vocalizations (CV) and near adult vocalizations (AV) were excluded, and 2-sec pauses were inserted when neither target

Table 3. Means and Standard Deviations for Analyzed Variables and Their Component Variables in Study 2

Variable	*M*	*SD*
RVC	.14	.05
Probability of chance occurrence of CV→AV→CV sequence	.0003	.0004
Probability of observed occurrence of CV→AV→CV sequence	.001	.001
Probability of CV	.05	.02
Probability of the CV→AV sequence	.006	.004
Probability of the AV→CV sequence	.03	.03
Consonant inventory	10.5	5

Note. CV = child vocalization. AV = adult vocalization. The symbol → = followed by. The consonant inventory score ranges from 0 to 20. RVC = reciprocal vocal contingency.

child nor near adult vocalizations occurred for at least 2 sec. When CV and AV overlapped, the segments were labeled by LENA software as a separate acoustic event (i.e., overlap) and excluded. Second, the new software program tallied three-event sequences into one of four cells in the 2 × 2 contingency table as described in Study 1 (Table 1 and Figure 1). The program output provided the counts of the key event types (i.e., CV and AV as labeled by LENA, as well as pause events that were inserted by the program) and the cell values for the 2 × 2 contingency table. RVC and the estimated chance probability of CV→AV→CV were then computed using the formulae provided in the description of Study 1. Finally, the observed probability of the CV→AV→CV sequence was computed by dividing the A cell count by the total number of CV, AV, and 2-sec pauses.

To aid future derivation of the RVC measure, we have made available free, cross-platform, open-source code for this software program [Yoder et al., 2016]. Step-by-step manuals for PC and Mac computers are also on the website. Number of events in the key sequence, key event types, and pause duration are user-defined variables in the program. Files can be processed in batches.

Measurement of speechlikeness of vocal communication. Speechlikeness of vocal communication was quantified as a consonant inventory in communication acts as manually coded from the Communication and Symbolic Behavior Scales-Developmental Profile Behavior Sample [CSBS-DP, Wetherby & Prizant, 2002]. The CSBS-DP is a standardized, structured communication sample designed to assess the communicative competence (i.e., use of eye gaze, gestures, vocalizations, words, understanding, and play) of children with a functional communication age between 6 and 24 months (chronological age approximately 6 months to 6 years in children with ASD). The CSBS-DP is widely used in research on children with ASD in the preverbal and early stages of language development.

The metric used in analyses was the weighted raw score from Subscale 11 (i.e., "Inventory of Consonants") from the CSBS-DP. Following the CSBS manual, the consonants considered were m, n, b, p, d, t, g, k, y, w, l, s, and sh. A consonant is counted as present if the child uses it in at least one communication act (i.e., vocalizations, symbols, or gestures directed to an adult with an apparent meaning). Voiced and unvoiced cognates (i.e., sounds produced in the same place of articulation) are not distinguished or credited separately in this consonant inventory. There are three pairs of cognates in the list (i.e., b and p, d and t, g, and k). Thus, the raw score for this measure had a minimum of 0 and a maximum of 10. According to standard CSBS-DP scoring, raw scores were weighted by 2, yielding a possible score range of 0–20. This variable has also been called diversity of key consonants used in communication [Woynaroski et al., 2016].

Interobserver reliability on the consonant inventory was estimated using a random sample of 20% of all coded sessions. Randomly selected reliability sessions were independently coded by a second observer, and primary coders were not aware which sessions would be selected for reliability checks. Reliability was estimated using a two-level random model and an absolute agreement intraclass correlation coefficient (ICC). We used a method of reliability estimation that accounted for (a) errors in identification of communication acts (unitization), (b) classification of whether the communication act included a consonant, and (c) identification of particular consonants in the same reliability estimate. The ICC for consonant inventory in communication acts was .98.

Results of Study 2

Preliminary results. Table 3 presents the means and standard deviations for the analyzed variables. Of particular interest, the average RVC was significantly greater than zero (i.e., chance), $t(20) = 11.9$, $P < .001$, and had a large one-sample effect size, Cohen's $d = 2.8$. This effect size indicates that the CV→AV→CV sequence occurred much more than expected by chance.

Short-term test-retest reliability of RVC. A one day interval was used between test and retest. The test-retest reliability was ICC = .64. When averaged across two daylong samples, the test-retest reliability increased to .78. The latter level of test-retest reliability is considered high [Yoder & Symons, 2010].

Correlation of RVC with consonant inventory. There was a strong positive correlation between RVC and consonant inventory, our measure of

Table 4. Zero-Order Correlations of RVC and an Alternative Possible Index of Child Vocal Reciprocity† with the Probabilities of Their Component Behaviors and Consonant Inventory in Study 2

	Probability of AV	Probability of CV	Probability of CV→AV sequence	Consonant inventory
RVC	−.13	.51**	.24	.60**
Probability of occurrence of CV→AV→CV sequence†	.53**	.75**	.86**	.43

Note. ** $P < .01$. AV = adult vocalization. CV = child vocalization. → = followed by. RVC = reciprocal vocal contingency.
†The alternative index does not control for chance sequencing of events.

speechlikeness of child vocal communication, $r = .60$, $P < .01$, 95% CI [.22, .82]. That is, as RVC increased, there was a tendency for children's consonant inventory to increase. Although the RVC score was designed to control for the chance probabilities of CV, AV, and various combinations thereof, RVC had an unexpected correlation with the probability of CV, $r = .51$, $P = .002$. We thus further tested the association of RVC and consonant inventory after statistically controlling for (a) the probability of CV, partial $r = .45$, $P = .045$, (b) the probability of AV, partial $r = .60$, $P = .005$, (c) the probability of the CV→AV sequence, partial $r = .58$, $P = .005$, and (e) the probability of AV→CV, partial $r = .67$, $P = .003$. Additionally, the association of RVC and speechlikeness of child vocal communication after controlling for the probability of the chance sequencing of the CV→AV→CV sequence was large, partial $r = .59$, $P = .004$. That is, the association of RVC with consonant inventory, our measure of speechlikeness of child vocal communication, did not occur due to covariation with these variables.

Post hoc analyses. To provide a basis of comparison to the RVC score, we examined the convergent validity of the observed probability of the three-event sequence CV→AV→CV, which differs from RVC in that it does not control for chance sequencing or the probabilities of events comprising the sequence. Table 4 presents the zero-order correlations for both this putative index of child vocal reciprocity and the RVC score with their component behavior probabilities and consonant inventory.

As the reader can see in Table 4, the zero-order correlation of the observed probability of the three-event sequence CV→AV→CV and consonant inventory was moderate in size, but statistically nonsignificant. The observed probability of the three-event sequence CV→AV→CV was furthermore strongly related to the probabilities of its component behaviors. Of particular importance, the observed probability of the CV→AV→CV sequence was highly correlated with the probability of CV. Consequently, the observed probability of the CV→AV→CV sequence was not correlated with consonant inventory after controlling for the probability of CV, $r = .12$, $P = .60$.

Logically, children may show good RVC regardless of where their chronological age or cognitive impairment level falls within the sample range. Additionally, children's RVC would not be expected to vary by parents' formal education. Thus, as evidence of divergent validation, we tested the correlations of the RVC with children's chronological age, $r = -.001$, IQ, $r = -.23$, and parents' formal education level, $r = -.27$. As expected, none were significantly related to RVC.

Discussion

Summary of Findings and Strengths of the Studies

In two studies, a new measure of child vocal reciprocity was proposed and tested: RVC, which we define as the positive sequential association of CV→AV with a subsequent CV. In Study 1, we confirmed the accuracy of the RVC score by showing that the mean RVC was almost identical to a known mean in a large number of simulated vocal samples. In Study 1, we also demonstrated that the method we used to estimate the frequency of chance sequencing of the three-event sequence of interest was accurate. The results provided the foundation for further vetting of RVC as a measure of child vocal reciprocity using real-life vocal samples (i.e., Study 2).

Study 2 provided psychometric information on the RVC score for a sample of preschoolers with ASD who were preverbal or in the early stages of language development. Results showed that the RVC index has high test-retest reliability when derived from two daylong vocal samples in children with ASD. For RVC to be useful as a correlate of children's vocal development, it must be temporally stable. Finally, we demonstrated that RVC was associated with speechlikeness of vocal communication in children with ASD with a large effect size.

We bolstered support for RVC as a psychometrically sound measure of child vocal reciprocity by showing that the correlation between children's RVC and the speechlikeness of their vocal communication was moderate to large even after statistically controlling for chance probability of CV→AV→CV or probabilities of the key sequence's component events. This result suggests that we succeeded in quantifying a dyadic construct—vocal reciprocity between a child and his or her adult communication partner/s—versus simply indexing the mere frequency of either child or adult vocalizations.

The aforementioned results stand in contrast to the findings for an alternative metric that we tested: the observed probability of the CV→AV→CV sequence (not controlling for chance occurrence). This latter measure might be considered by some as a potential measure of child vocal reciprocity; however, the observed probability of the CV→AV→CV sequence ceased to be associated with speechlikeness of vocal communication when the probabilities of its component behaviors were controlled. One of the component behaviors was CV. A reasonable interpretation of these results is that the observed probability of the CV→AV→CV sequence primarily reflects frequency of child vocalizations, rather than a dyadic exchange between a child and his or her adult communication partners.

We furthermore provided support for the divergent validity of RVC. As expected, RVC was not associated with other parent and child variables that were anticipated to be unrelated to child vocal reciprocity. Together, the findings of Study 2 begin the process of building the construct validity of our novel index of RVC in children with ASD.

Limitations

The primary limitation of Study 1 is that, like all simulation studies, the method of generating the data requires that the investigator make certain assumptions, which might not actually apply to real-life vocal samples of children with ASD. For example, the mean probability of CV and of AV were higher in the simulated vocal samples than in the real vocal samples of children with ASD. Additionally, the simulation indicated RVC was accurate for a known mean sequential association that was zero. Some readers might object to using simulated vocal samples with random sequencing of CV and AV because we would anticipate a positive RVC for most vocal samples of children in the early stages of language development. To our knowledge, though, simulations with a known mean sequential association are the only way to estimate the absolute accuracy of an index of sequential association. At present, it is only by randomly sequencing events in a very large number of samples that we can know the mean sequential association.

The limitations of Study 2 were its (a) small sample size, (b) concurrent correlational research design, and (c) inability to distinguish among adults in the daylong vocal samples. Sampling theory tells us that a small sample size reduces the probability of replication relative to findings from larger samples. The concurrent correlational design precludes inferences regarding direction of effect as well as causal influence, as it does not establish temporal precedence or control for third variable explanations of observed associations. For example, it is possible that variability in children's receptive language has a causal influence on both RVC and speechlikeness of vocal communication. Finally, because the software analyzing the daylong vocal samples does not distinguish among different adults, we cannot interpret the RVC as being specific to a particular child-adult dyad. Instead, we propose that the RVC is a generalized measure of a child's attention and response to preceding vocal responses from a broad range of adult communication partners.

Future Research

Additional work is needed to test the value and limitations of the RVC as a measure of vocal reciprocity in children on the autism spectrum. Future work is needed to determine how context and time of sampling affects the validity of the RVC. Longitudinal correlational studies that establish temporal precedence of RVC relative to speechlikeness will increase our confidence that early vocal reciprocity is useful for predicting future speechlikeness of vocal productions in preschoolers with ASD. Longitudinal studies will additionally be helpful for determining the extent to which this index of vocal reciprocity is sensitive to change over the course of development. Experimental studies that control for possible alternative explanations for the associations we have observed will lend important insights into the causal nature of these relations. Well-controlled clinical trials will be particularly informative regarding whether the RVC score is sensitive to effects of treatment, and whether early effects of intervention on RVC scores may precede and mediate effects of treatment on children's speechlikeness. Future research involving a predictive association of early RVC to later measures of child semantically-related verbal response to prior adult response to child behavior using a method that controls for chance adjacency of behaviors will potentially add to the validity evidence supporting RVC. One might consider such an association evidence of criterion-related validity if the later measure gains gold standard status eventually.

Future research involving larger samples of children with ASD in varying stages of language development is also necessary to test whether RVC's predictive convergent and divergent validity, as well as sensitivity to change, vary as a function of the initial language level of children on the autism spectrum. Of particular interest is the prediction of future expressive language ability in this clinical population. We suspect that child vocal reciprocity should be predictive of not only speechlikeness of vocal communication, but also broader language development. The process by which child vocal reciprocity affects broader expressive language acquisition is likely to unfold over several months or even years in some children with ASD though [Warlaumont et al., 2014]. If the predictive relation with expressive language is supported, then testing the malleability of RVC in response to intervention will be warranted. If RVC is malleable, it could represent a useful early intervention target for young, preverbal children with ASD.

Potential Developmental Implications

If future research continues to support the validity of the RVC, then RVC could inform our understanding of the ways that child and adult factors influence the development of language in children with ASD. As adults are generally more responsive to children with ASD than children with ASD are responsive to adults, growth in RVC could show growth in children's attention to, processing of, and response to adult vocal responses to children's vocalizations. If so, then children with ASD who show faster growth in RVC or who have achieved relatively high levels of RVC may be more likely to benefit from the types of adult linguistic input that tends to facilitate language development. Although the current study's concurrent design and developmentally young sample prevented showing that RVC is associated with broader language development, we suspect that RVC should predict later language level, at least in the early stages of language learning. As we suggested above, this is an important area for future research.

Additionally, RVC could be a moderator of the association between adult linguistic input and child language development in children with ASD. Past work showed that adult linguistic mapping of immediately preceding child intentional communication was predictive of later language only in children with ASD at or above a certain level of receptive language [McDuffie & Yoder, 2010]. As RVC may predict future language, it is possible that RVC will be an earlier-occurring moderator of the association between adult linguistic responses and later language learning.

Possible Implications for the Study of Reciprocity in Children with ASD

The RVC score is focused on one important aspect of the reciprocity deficit in children with ASD—vocal reciprocity. It is likely to be important that the RVC is tracking children's vocal responses, not to adult vocalizations that occur in isolation, but rather to adults' vocal responses to children's vocalizations. Prior work has shown that children with ASD are more likely to attend to and benefit from adult linguistic responses to children's communication than to adult linguistic input that is not in response to children's behavior [McDuffie & Yoder, 2010]. Past work operationalizing "reciprocity" as a two-event sequence (e.g., research on conversational turn count metrics) assumes that vocal reciprocity can be quantified through child responses to adult behaviors without attention to whether the adult behavior is a response to child behavior. Such work, in our opinion, evaluates "responsiveness" rather than "reciprocity"—two related, but importantly different concepts.

Clinical Implications

If future research continues to support the validity of the RVC in children with ASD in the early stages of language development, then clinicians could utilize the LENA recording devices to collect two daylong vocal samples, then use the LENA Pro and provided software to efficiently measure child vocal reciprocity in their everyday clinical practice. If our hypotheses about the RVC score are born out in future work, this index may be useful for treatment planning, as well as for progress monitoring in interventions targeting vocal reciprocity, speechlikeness of vocal communication, and broader expressive language acquisition in children with ASD.

Summary

We have introduced a new measure of child vocal reciprocity—RVC. The two studies presented in this manuscript found initial evidence that the new measure is accurate, reliable and construct valid in children with ASD who are preverbal or in the early stages of language development. This index of vocal reciprocity can be automatically derived from daylong vocal samples collected in children's natural environment using a new, freely available software program. We have provided a URL in the reference section [Yoder et al., 2016] where future users can access the free software program and user manuals. We hope that others will capitalize on this resource to advance our understanding of the scientific utility (and potential limitations) of the RVC score.

Acknowledgments

This research was supported by the National Institute for Deafness and other Communication Disorders (NIDCD R01 DC006893); the Office of Special Education Programs, U. S. Department of Education (H325D100010); the National Science Foundation under grant numbers BCS-1529127 and SMA-1539129; the National Center for Advancing Translational Sciences (KL2TR000446); and the *EKS* NICHD (U54HD083211). Its contents are solely

the responsibility of the authors and do not necessarily represent the official views of the National Center for Advancing Translational Sciences, the National Science Foundation, the National Institutes of Health, or the U. S. Department of Education. We derive no financial or non-financial benefits from the content or software program discussed in this manuscript.

References

American Psychiatric Association. (2000). Diagnostic and statistical manual of mental disorders (4th ed., text rev. ed.). Washington, DC: Author.

Bakeman, R., & Gottman, J.M. (1997). Observing interaction: An introduction to sequential analysis. Cambridge: Cambridge University Press.

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

Brundin, K., Rödholm, M., & Larsson, K. (1988). Vocal communication between parents and infants. Early Human Development, 16, 35–53.

Crocker, L., & Algina, J. (1986). Introduction to classical and modern test theory. Fort Worth, TX: Harcourt.

Cronbach, L.J., & Meehl, P.E. (1955). Construct validity in psychological tests. Psychological Bulletin, 52, 281.

Dunst, C.J., Hamby, D.W., Trivette, C.M., Prior, J., & Derryberry, G. (2013). Effects of a socially interactive robot on the conversational turns between parents and their young children with autism. Social Robots Research Reports, Number 6. Orelena Hawks Puckett Institute.

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

Ganek, H.V., & Eriks-Brophy, A. (2017). A concise protocol for the validation of Language ENvironment Analysis (LENA) conversational turn counts in Vietnamese. Communication Disorders Quarterly, 39, 371–380.

Goldstein, M.H., Waterfall, H.R., Lotem, A., Halpern, J.Y., Schwade, J.A., Onnis, L., & Edelman, S. (2010). General cognitive principles for learning structure in time and space. Trends in Cognitive Sciences, 14, 249–258.

Gottman, J.M., & Roy, A.K. (1990). Sequential analysis: A guide for behavioral researchers. Cambridge: Cambridge University Press.

Gros-Louis, J., West, M.J., Goldstein, M.H., & King, A.P. (2006). Mothers provide differential feedback to infants' prelinguistic sounds. International Journal of Behavioral Development, 30, 509–516.

Gros-Louis, J., West, M.J., & King, A.P. (2014). Maternal responsiveness and the development of directed vocalizing in social interactions. Infancy, 19, 385–408.

Hammond, L.J. (1980). The effect of contingency upon the appetitive conditioning of free-operant behavior. Journal of the Experimental Analysis of Behavior, 34, 297–304.

Klin, A. (1991). Young autistic children's listening preferences in regard to speech: A possible characterization of the symptom of social withdrawal. Journal of Autism and Developmental Disorders, 21, 29–42.

Kuhl, P.K. (2003). Human speech and birdsong: Communication and the social brain. Proceedings of the National Academy of Sciences, 100, 9645–9646.

LENA Research Foundation. (2016). LENA Pro. Retrieved from http://www.lenafoundation.org/lena-pro

Lloyd, B.P., Kennedy, C.H., & Yoder, P.J. (2013). Quantifying contingent relations from direct observation data: Transitional probability comparisons versus Yule's Q. Journal of Applied Behavior Analysis, 46, 479–497. doi:10.1002/jaba.45

Lloyd, B.P., Yoder, P.J., Tapp, J., & Staubitz, J.L. (2015). The relative accuracy and interpretability of five sequential analysis methods: A simulation study. Behavior Research Methods, 48, 1482–1491.

Lord, C. (1995). Follow-up of two-year-olds referred for possible autism. Journal of Child Psychology & Psychiatry, 36, 1365–1382.

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, 205–223.

Martens, B.K., Gertz, L.E., Werder, C.S., Rymanowski, J.L., & Shankar, K.H. (2014). Measures of association in contingency space analysis. Journal of Mathematical Psychology, 59, 114–119. doi:10.1016/j.jmp.2013.06.004

McDuffie, A., & Yoder, P. (2010). Types of parent verbal responsiveness that predict language in young children with autism spectrum disorder. Journal of Speech, Language, and Hearing Research, 53, 1026–1039.

Mullen, E.M. (1995). Mullen scales of early learning (AGS ed.). Los Angeles: Western Psychological Services.

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, 523–547.

Patten, E., Belardi, K., Baranek, G.T., Watson, L.R., Labban, J.D., & Oller, D.K. (2014). Vocal patterns in infants with autism spectrum disorder: Canonical babbling status and vocalization frequency. Journal of Autism and Developmental Disorders, 44, 2413–2428.

Paul, R., Chawarska, K., Cicchetti, D., & Volkmar, F. (2008). Language outcomes of toddlers with autism spectrum disorders: A two year follow-up. Autism Research, 1, 97–107.

Shriberg, L.D., Paul, R., Black, L.M., & van Santen, J.P. (2011). The hypothesis of apraxia of speech in children with autism spectrum disorder. Journal of Autism and Developmental Disorders, 41, 405–426.

Staubitz, J.L., & Lloyd, B.P. (2016, May). Evaluating the generalizability of direct measures of problem behavior and contingencies in elementary general education classrooms. Poster session presented at the annual meeting of the Society for Quantitative Analyses of Behavior, Chicago, IL.

Warlaumont, A.S., Richards, J.A., Gilkerson, J., & Oller, D.K. (2014). A social feedback loop for speech development and its reduction in autism. Psychological Science, 25, 1314–1324.

Wetherby, A., & Prizant, B. (2002). Communication and symbolic behavior scales developmental profile—First normed edition. Baltimore, MD: Paul H. Brookes.

Woynaroski, T., Oller, D.K., Keceli-Kaysili, B., Xu, D., Richards, J.A., Gilkerson, J., Gray, S., & Yoder, P. (2017). The stability and validity of automated vocal analysis in preverbal preschoolers with autism spectrum disorder. Autism Research, 10, 508–519. doi:10.1002/aur.1667

Woynaroski, T., Watson, L., Gardner, E., Newsom, C.R., Keceli-Kaysili, B., & Yoder, P.J. (2016). Early predictors of growth in diversity of key consonants used in communication in initially preverbal children with autism spectrum disorder. Journal of Autism and Developmental Disorders, 46, 1013–1024.

Xu, D., Yapanel, U., Gray, S., & Baer, C.T. (2008). The LENA language environment analysis system: The interpreted time segments file. Retrieved from https://www.lenafoundation.org/wp-content/uploads/2014/10/LTR-04-2_ITS_File.pdf

Yoder, P., & Symons, F. (2010). Observational measurement of behavior. New York: Springer.

Yoder, P., Wade, J., Tapp, J., Warlaumont, A., & Harbison, A.L. (2016). Contingencies from LENA data. Retrieved from https://github.com/HomeBankCode/LENA_contingencies

Yoder, P., Watson, L.R., & Lambert, W. (2015). Value-added predictors of expressive and receptive language growth in initially nonverbal preschoolers with autism spectrum disorders. Journal of Autism and Developmental Disorders, 45, 1254–1270.