



Validity of Vocal Communication and Vocal Complexity in Young Children with Autism Spectrum Disorder

Jena McDaniel¹ · Paul Yoder² · Annette Estes³ · Sally J. Rogers⁴

Published online: 9 October 2019

© Springer Science+Business Media, LLC, part of Springer Nature 2019

Abstract

To identify valid measures of vocal development in young children with autism spectrum disorder in the early stages of language learning, we evaluated the convergent validity, divergent validity, and sensitivity to change (across 12 months) of two measures of vocal communication and two measures of vocal complexity through conventional coding of communication samples. Participants included 87 children with autism spectrum disorder ($M = 23.42$ months at entry). All four vocal variables demonstrated consistent evidence of convergent validity, divergent validity, and sensitivity to change with large effect sizes for convergent validity and sensitivity to change. The results highlight the value of measuring vocal communication and vocal complexity in future studies.

Keywords Autism spectrum disorder · Language · Validity · Vocalizations

Vocal development in children with autism spectrum disorder (ASD) is an understudied area with potential clinical utility for enhancing language trajectories. Improving language trajectories and language outcomes for children with ASD is critical because language skills are affected by ASD and language competence predicts social, adaptive, and vocational outcomes in this population (Billstedt et al. 2005; Howlin 2000). For children with ASD, assessing and targeting vocal development, the process through which children produce increasingly speech-like sounds (Oller 2000), may be useful for three reasons. First, compared with only targeting lexical or grammatical development, targeting

vocal communication during the preverbal stage of communication development might be more effective in facilitating language development in children who are not ready for linguistic targets. Second, because vocal development is a logical precursor of language, vocal measures might show early response to intervention that targets lexical development and help explain why intervention initially targeting vocalizations is effective in facilitating language in preverbal children with ASD (e.g., Woynaroski et al. 2014; Yoder et al. 2015b). Finally, evaluating vocal development prior to initiating communication intervention may provide a means for determining which children are more likely to benefit from a given intervention approach (Yoder and Warren 2002). To experimentally evaluate these reasons for focusing on vocal development, researchers must employ valid vocal development measures.

Because there is no gold-standard vocal development measure, one cannot simply correlate a new vocal measure with a gold-standard vocal development measure to evaluate the new measure's validity (i.e., criterion-related validity). Instead, one must draw on multiple sources of evidence to assess the degree to which a variable demonstrates that it measures what it purports to measure (i.e., construct validation; Cronbach and Meehl 1955). Construct validity includes convergent validity (i.e., degree to which a variable correlates with other variables with which it is predicted to correlate based on theory) and divergent validity (i.e., a variable does *not* correlate other

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s10803-019-04248-x>) contains supplementary material, which is available to authorized users.

✉ Jena McDaniel
jena.c.mcdaniel@ku.edu

¹ Schiefelbusch Institute for Life Span Studies, University of Kansas, 3001 Dole Human Development Center, 1000 Sunnyside Ave, Lawrence, KS 66045-7555, USA

² Department of Special Education, Vanderbilt University, Nashville, USA

³ Department of Speech and Hearing Sciences, University of Washington, Seattle, USA

⁴ Department of Psychiatry and Behavioral Sciences, University of California, Davis, Davis, USA

variables with which it is not predicted to correlate based on theory; Campbell and Fiske 1959). Sensitivity to change, the degree to which a measure changes over time, is another key feature of high-quality measures, particularly those used to examine change occurring within intervention studies. This study assesses the construct validity and sensitivity to change of four variables purported to capture vocal communication or vocal complexity in children with ASD.

Two Potentially Important Aspects of Vocalizations for Children with ASD

Vocal communication is defined as how frequently or consistently a child produces vocalizations directed towards another person in an apparent attempt to transmit a message (Wetherby et al. 1989). These vocalizations appear to be used to initiate or maintain a social interaction. Because children do not always direct their vocalizations to another person, one would expect fewer communicative vocalizations than total vocalizations.

Vocal complexity is defined as the frequency, consistency, or diversity with which a child produces vocalizations with speech-like features, such as canonical syllables or consonants (e.g., Wetherby et al. 2007; Woynaroski et al. 2017; Yoder et al. 2015a). Typically developing children progress from producing quasivowels (0–2 months), gooing (1–4 months), grunts, squeals, fully resonant vowels, and marginal babbling (3–8 months) to canonical babbling (5–10 months; Oller 2000). Canonical babbling sounds substantially more like adult speech than precanonical vocalizations because it includes vowel-like and consonant-like sounds with rapid, adult-like transitions (Oller et al. 1999). Example vocal complexity measures include the rate of consonant–vowel productions (Talbot 2014), proportion of vocalizations with a canonical syllable (Woynaroski et al. 2017), and diversity of key consonants used in communication acts (DKCC; Wetherby et al. 2007; Woynaroski et al. 2017). DKCC is calculated as how many of 13 consonants (i.e., /m/, /n/, /b/ or /p/, /d/ or /t/, /g/ or /k/, /w/, /l/, “j,” /s/, and “sh”) that a child produces during the sample. Children can only receive up to ten points because members of voiced-voiceless pairs (e.g., /b/ vs. /p/, /d/ vs. /t/, /g/ vs. /k/) are difficult to distinguish reliably on recordings. Thus, children can only receive one point for the pair regardless of whether they produce one or both of each pair’s consonants.

Theoretical Support for Measuring Vocal Communication and Vocal Complexity

Before assessing the construct validity and sensitivity to change of vocal variables, one must consider the theoretical basis for variables to which vocal variables should

correlate (i.e., convergent validity), for variables to which vocal variables should not correlate (i.e., divergent validity), and changes in the vocal variables over time (i.e., sensitivity to change). Child-driven theories of language development focus on child characteristics, such as those related to speech production, and assert that vocal communication and vocal complexity facilitate language development in children with ASD. These child-driven explanations include shared articulators (e.g., tongue and lips) and shared motor pathways for speech and vocalizations (Fry 1966; Iverson 2010; Stoel-Gammon 2011; Vihman 1992, 1996) for prelinguistic vocalizations and spoken words. For instance, regardless of whether the production of “wawa” is produced without lexical meaning or as an approximation of “water,” the articulators move in the same way. One can also make theoretical arguments involving bidirectional influence between child characteristics and adult input, rather than focusing only on child characteristics.

Social Feedback Theory

Goldstein et al. (2003) and Goldstein and Schwade (2008) presented the social feedback theory as a potential explanation for vocal development in infants with typical development. This theory emphasizes the role of contingent caregiver responses to child vocalizations in social interactions for facilitating more complex child vocalizations over time. Goldstein et al. (2003) asserted that infants produce more complex and more adult-like vocalizations following contingent adult responses within social interactions compared with noncontingent adult responses. Based on results of an experimental study with 6- to 10-month-old infants, Goldstein and Schwade (2008) concluded that infants produced either more fully resonant vowels or more consonant–vowel syllables depending on how caregivers contingently responded to the children’s vocalizations (i.e., with a fully resonant vowel or a consonant–vowel syllable, respectively). However, results of the key comparison between the contingent response group and the corresponding control group were not reported, even though this comparison was possible with the study design. Therefore, direct evidence of the impact of contingent versus noncontingent responses was not provided.

Social Feedback Loop

Warlaumont et al. (2014) proposed a “social feedback loop” in which (a) adults are more likely to respond to child vocalizations that are speech-related and (b) a child is more likely to produce speech-related vocalizations if an adult responded immediately to the child’s preceding utterance. Speech-related vocalizations include words as well as prespeech vocalizations (i.e., babbling; Oller et al.

2010). They posited that this social feedback loop may be disrupted in children with ASD because (a) children with ASD produce fewer speech-related vocalizations than children with typical development, (b) caregivers of children with ASD respond differently than caregivers of children with typical development, and (c) children with ASD have a reduced ability to respond to adults' contingent responses. This disruption might explain in part the language deficits in children with ASD.

Transactional Theory of Spoken Language Development

The transactional theory of spoken language development considers child factors (e.g., cognitive, social, and motor abilities), parent factors (e.g., linguistic input), and dyadic factors (i.e., parent–child) while emphasizing the bidirectional nature of the interactions between child and parent factors across development (Camarata and Yoder 2002; McLean and Snyder-McLean 1978; Sameroff and Chandler 1975; Woynaroski et al. 2014). It posits that as a child's speech and language skills increase, parents provide more complex input that scaffolds continued child growth (Camarata and Yoder 2002; Woynaroski et al. 2014). Despite some mixed findings, bidirectional influences have been documented for vocal development in children with typical development (Fagan and Doveikis 2017). For example, the content of infant vocalizations and accompanying actions (e.g., play actions and directing eye gaze) has been reported to influence how mothers respond to infant vocalizations (West and Rheingold 1978; Yoder and Feagans 1988). In initially preverbal children with ASD, both parent factors (i.e., linguistic input) and child factors (i.e., intentional communication and receptive vocabulary) predicted growth in DKCC, which in turn predicted expressive language in children with ASD (Woynaroski et al. 2017; Yoder et al. 2015a). These findings are consistent with the transactional theory of spoken language development in children with ASD. Positive associations between parent verbal responsiveness (e.g., follow-in comments) to child behaviors and child spoken language skills in children with ASD provide additional support (e.g., McDuffie and Yoder 2010). In contrast to these findings supportive of bidirectional influences, Fagan and Doveikis (2017) reported that mothers responded to a relatively small number of infant vocal behaviors (i.e., 30%) during ordinary interactions. However, mothers responded much more to infant vocalizations than other infant vocal behaviors (e.g., crying, coughing, and raspberries). Thus, the difference in types of infant vocal behaviors coded may explain the differences in findings.

Application to Vocal Development in Children with ASD

Theoretically, improving vocal communication and/or vocal complexity could facilitate language development in children with ASD. The social feedback theory, social feedback loop, and transactional theory of spoken language development all emphasize interactions between adults and children in children's vocal development. Increasing vocal communication or vocal complexity in children might elicit more frequent or complex adult responses to scaffold the child's ability to produce more adult-like productions including spoken words. Also, increases in vocal communication or vocal complexity could signal that children are attempting to say words they understand but do not yet produce accurately enough to be understood (Woynaroski et al. 2016).

Current Empirical Support for Measuring Selected Aspects of Vocalizations

When establishing the convergent validity of particular vocal variables for specific purposes, correlations between vocal variables of interest with expressive language outcomes or measures of precursors to expressive language are some of the most relevant pieces of evidence. Evidence from studies that use a longitudinal correlational design, rather than a concurrent correlational design, and that include children with ASD in the early stages of language learning, are most relevant here. Longitudinal associations provide stronger evidence of convergent validity than concurrent associations because they document an association between variables and provide evidence of a temporal precedence of the putative cause relative to the putative effect.

Broadly, a recent meta-analysis revealed that vocalizations correlate strongly with current or future expressive language skills for children with ASD (McDaniel et al. 2018). The meta-analysis included a variety of vocal variables, including those purported to measure vocal communication and vocal complexity. However, the overall number of studies was too low to achieve sufficient power to test whether specific types of variables yielded stronger correlations with expressive language than others.

Vocal Communication

There is equivocal evidence regarding the correlations between communicative vocalizations specifically and expressive language in children with ASD. For example, Plumb and Wetherby (2013) reported that the proportion of vocalizations that were communicative in the second year of life predicted expressive language skills at age 3 above and beyond the proportion of vocalizations that were

noncommunicative. In contrast, for children suspected of having ASD (mean chronological age = 19.51 months), Swineford (2011) reported nonsignificant correlations between the rate of communication acts with vocalizations and verbal communication concurrently and predictively.

Vocal Complexity

Vocal complexity has been defined in multiple ways within two broad categories: (a) vocalizations with consonants and/or canonical syllables without differentiating diversity of consonants produced and (b) diversity of consonants produced. Within each of these two broad categories, there are two subordinate categories: those that are derived from (a) all vocalizations versus (b) only communicative vocalizations.

When only communicative vocalizations are used, the vocal complexity variable combines communication and complexity concepts. For example, several studies have examined consonant inventories in communicative vocalizations of children with ASD, rather than in all vocalizations. For the analyses and discussion purposes, we classify these variables within the complexity set because we judged the complexity component to be more prominent in the variable's interpretation than the communication component. For example, DKCC is conceptually more related to consonant inventory in all vocalizations, which is clearly a complexity variable, than to the number of vocal communication acts. Relatedly, proportion of communication acts with a canonical syllable is judged to be a vocal complexity variable because it focuses on how consistently the child uses canonical syllables (a marker of complexity) rather than how consistently a child uses vocalizations for communicative purposes.

There is empirical evidence for correlations between the types of vocal complexity variables described above and expressive language in children with ASD. For use of vocalizations with consonants, the rate of consonant–vowel vocalizations produced at 9 months of age correlated with expressive language at 12 months of age for children with ASD (Talbot 2014). Degree of verbal delay concurrently and negatively correlated with lack of communicative vocalizations with consonants (Book 2009; McCoy 2013). Relatedly, the rate of canonical babbling correlated with concurrent expressive language in children with ASD (mean chronological age = 44.67 months; Sheinkopf et al. 2000).

For consonant inventory measures, Yoder et al. (2015a) found that DKCC predicted expressive language growth in initially preverbal children with ASD over and above ten other putative predictors. Similarly, Wetherby et al. (2007) identified that DKCC at 18–24 months was one of the “best predictors of verbal skills at 3 years” (p. 971), compared with numerous other possible predictors for children with

ASD. Relatedly, a composite variable derived from the proportion of communication acts with a canonical syllable and DKCC strongly correlated with later expressive vocabulary in a sample of initially preverbal children with ASD (Woy-naroski et al. 2017).

The Need for Additional Evidence of Validity for Vocal Variables

A single analysis or test is insufficient for reporting the degree to which a measure exhibits construct validity (Cronbach and Meehl 1955). Instead, multiple sources of evidence must be integrated and evaluated for the specific purpose of the variable of interest. When comparing evidence of validity among multiple vocal variables, one can compare variables with attention to the number of different purposes the variable might serve. For example, a measure may exhibit strong evidence of convergent construct validity with expressive language, but weak evidence that it is sensitive to change. Currently, nearly all validity evidence for measuring vocal development in young children with ASD is convergent validity evidence, even though divergent validity is a key type of validity evidence. The current study presents an opportunity to compare directly convergent validity, divergent validity, and sensitivity to change of selected vocal variables from the same, relatively large sample of young children with ASD. It begins to fill this gap in the literature and move the field forward in selecting vocal measures that are most likely to yield meaningful, interpretable results.

For evidence of convergent validity, we test whether vocal variables predict expressive language, which is predicted by theory and evidence (e.g., Goldstein et al. 2003; Goldstein and Schwade 2008; McDaniel et al. 2018; McLean and Snyder-McLean 1978; Sameroff and Chandler 1975; Warlaumont et al. 2014). For evidence of divergent validity, we test whether vocal variables predict nonverbal cognitive skills, which they would *not* be expected based on theory. Cognition and language are separate constructs. Particularly in children with disabilities, such as ASD, individuals can present with relatively low expressive language skills in the context of average or even above average nonverbal cognitive skills (e.g., Stark and Tallal 1981). Although no known studies provide evidence of divergent validity for the included vocal variables, evidence that vocal variables do not predict nonverbal cognitive skills would support an inference that the vocal variable is specific to vocal development rather than a more general developmental measure.

For evidence of sensitivity to change, we examine whether vocal variables exhibit a change in value from study initiation to 12 months later. The ability to capture change is a particularly important feature for an intervention study's outcome measure. The social feedback theory,

social feedback loop, and transactional theory of spoken language development assert that vocalizations are expected to increase in complexity and to be used for communicative purposes as children develop (Goldstein et al. 2003; Goldstein and Schwade 2008; McLean and Snyder-McLean 1978; Sameroff and Chandler 1975; Warlaumont et al. 2014). Therefore, vocal variables assessing complexity and communicative use are expected to increase over time.

Purpose and Research Questions

This study evaluates and compares quantitative evidence of validity for vocal variables purported to assess vocal development in young children with ASD. We addressed the following research questions for two vocal variables purported to assess vocal communication (i.e., number of communication acts that include a vocalization and proportion of vocalizations that are communicative) and two purported to assess vocal complexity (i.e., DKCC and proportion of vocalizations with a canonical syllable). (1) To assess convergent validity, does the vocal variable predict later expressive language skills? (2) To assess divergent validity, does the vocal variable *not* predict later nonverbal cognitive skills? (3) Does the vocal variable exhibit sensitivity to change?

Method

Participants

The study includes 87 children (21 female, 66 male) who participated in the Toddlers with Autism: Developing Opportunities for Learning (TADPOLE) multi-site randomized controlled trial (Rogers et al. 2013). The TADPOLE study compared language and developmental outcomes of a sample of young children with ASD who were randomly assigned to one of two treatment styles and one of two treatment intensity levels. However, the results used for evidence of construct validity and sensitivity to change were not influenced by groups to which participants were assigned as indicated by nonsignificant predictor by group interactions. When an interaction term for style by predictor or intensity by predictor was added to each growth curve model used in the convergent validity analyses, the interaction term was nonsignificant. Similarly, for the sensitivity to change analyses, when style by time or intensity by time interaction terms were added, none were significant.

Participants met the following inclusion criteria: (a) chronological age of 13 to 30 months at study entry, (b) ambulatory without primary motor impairments affecting hand use, (c) meets ASD diagnostic criteria, (d) overall developmental quotient of at least 35 on the Mullen Scales

of Early Learning (MSEL; Mullen 1995), (e) English as a primary language (i.e., English reportedly spoken at least 60% of the time at home), and (f) hearing and visual acuity within normal limits per screening. ASD diagnosis was based on: (a) Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition diagnostic criteria for ASD (American Psychiatric Association 2013), (b) clinical consensus of diagnosis based on record review and observation by two independent staff, one of whom is a licensed psychologist, (c) meeting full autism criteria on the Autism Diagnostic Interview-Revised (Lord et al. 1994), (d) meeting autism cutoff on the Autism Diagnostic Observation Schedule for Toddlers (Luyster et al. 2009), and (e) diagnosis confidence rating of relatively confident or very confident assigned by the assessor who evaluated the child. Participants were not excluded based on the presence of genetic disorders or other health conditions.

Per caregiver report, 48 participants were reported to be white, 19 to be more than one race, 9 to be Asian, 7 to be black or African American, 1 to be American Indian/Alaskan native, 1 to be Native Hawaiian or other Pacific Islander, and 2 as unknown. Seventeen participants were reported to be Hispanic/Latino, 64 to be non-Hispanic, and 6 as unknown. Maternal education level was reported as follows: 1 had some high school, 6 had a high school diploma, 25 had some college, 24 had a college degree, 6 had some graduate school, 22 had a graduate degree, and 1 reported “other.” See Table 1 for additional participant characteristics.

Procedures

The study’s constructs, procedures, and variables are listed in Table 2. Data are used from procedures administered across three time periods that spanned 12 months (Time 1 = study initiation; Time 2 = 6 months post study initiation; Time 3 = 12 months post study initiation).

Communication Sample Procedure

The Communication Sample Procedure (CSP) is a 15-min semi-structured free-play communication sample with a standard toy set in a lab setting with interspersed opportunities for the child to request clarification and to respond to an examiner’s topic change. The examiner’s interaction style is guided by specific principles designed to support productive engagement (e.g., follow the child’s lead and join in and play at the child’s demonstrated level of play) and communication (e.g., talking about topics related to child’s focus of attention, monitoring utterance length and complexity, and avoiding directives) as described in the procedure manual. This manual is available upon request from the first author. The CSP was administered at Times 1 and 3.

Table 1 Participant characteristics at study entry

	Mean	SD	Min	Max
Chronological age (months)	23.42	3.98	13.78	30.71
Developmental quotient	58.83	17.96	31.04	121.98
MSEL receptive language (age equivalent in months)	10.11	7.22	1	33
MSEL expressive language (age equivalent in months)	11.97	4.71	4	27

Developmental quotient mean of age equivalent scores for fine motor, visual reception, receptive language, and expressive language on the Mullen Scales of Early Learning divided by chronological age multiplied by 100; *Max* maximum; *Min* minimum; *MSEL* Mullen Scales of Early Learning (Mullen 1995)

Early Communication Index (Greenwood et al. 2006; Luze et al. 2001)

The Early Communication Index (ECI), one of the Individual Growth and Development Inventories, is a 6-min play-based measure that uses a standard toy set in a lab setting. The ECI has been validated on multiple samples of young children that include more than 7000 total children (Greenwood et al. 2006, 2010; Luze et al. 2001). The samples include racially diverse children, most of whom attended Early Head Start and some of whom had a disability. The general principles of examiner behavior and talk followed in the CSP are followed in the ECI as well. The ECI may be used frequently to monitor progress during intervention. It was administered monthly throughout the 12-month period for a total of 13 administrations. To align with the data from the 15-min CSP, we averaged the ECI sessions from the first 3 months for Time 1 for a total of 18 min of coded time and the sessions from the last 3 months for Time 3 for a total of

18 min of coded time. Using three ECI sessions for each time point increased the stability of the coded variables.

MacArthur-Bates Communicative Development Inventory (Fenson et al. 2007)

Caregivers completed a compilation form (i.e., 720 total items from the Words and Gestures and Words and Sentences vocabulary items) of the MacArthur-Bates Communicative Development Inventory (MB-CDI) for expressive vocabulary at all three time points. Caregivers marked words on the checklist that they observed their child saying at least once in the prior 2 weeks.

Mullen Scales of Early Learning (Mullen 1995)

The Mullen Scales of Early Learning (MSEL) was administered at three time points. The MSEL includes subscale scores for receptive language, expressive language, visual reception, and fine motor skills.

Vineland Adaptive Behavior Scales, Second Edition (Sparrow et al. 2005)

The examiner interviewed the participants' caregiver(s) to complete the Vineland Adaptive Behavior Scales, Second Edition (VABS) at three time points. The VABS includes subscale scores for expressive language, daily living, and fine motor skills.

Coding Vocal Variables

Trained research assistants and the first author completed observational coding for the CSP and ECI using ProCoder

Table 2 Study constructs, procedures, and variables

Construct	Procedure (s)	Variable
Vocal communication	CSP & ECI	Number of communication acts that include a vocalization Proportion of vocalizations that are communicative
Vocal complexity	CSP & ECI	Diversity of key consonants used in communication acts Proportion of vocalizations with a canonical syllable
Expressive language skills	MB-CDI	Raw score for words said
	MSEL	Expressive subscale age-equivalency score
	VABS	Communication domain expressive subscale raw score
	CSP	Number of different word roots said
Nonverbal cognitive skills	VABS	Daily living skills subscale age-equivalency score Fine motor skills subscale age-equivalency score
	MSEL	Fine motor subscale age-equivalency score
		Visual reception subscale age-equivalency score

CSP Communication Sample Procedure; *ECI* Early Communication Index (Greenwood et al. 2006; Luze et al. 2001); *MB-CDI* MacArthur-Bates Communicative Development Inventory: Expressive vocabulary compilation form (Fenson et al. 2007); *MSEL* Mullen Scales of Early Learning (Mullen 1995); *VABS* Vineland Adaptive Behavior Scales, Second Edition (Sparrow et al. 2005)

DV (Tapp 2003) and Systematic Analysis of Language Transcripts (SALT) software (Miller and Chapman 2016). Coders completed four passes using timed-event behavior sampling to code behaviors in the CSP and ECI necessary for deriving the vocal development variables. On the first pass, the coder identified codable and uncodable portions of each video file. On the second pass, the coder identified all communication acts within the codable time and orthographically transcribed words that the child said. The coding manual, which is available from the first author, includes detailed communication act coding rules. See Table 3 for operational definitions of key concepts for coding. Within the third pass, the coder identified and classified vocalizations that occurred within a communication act to indicate whether they contained one or more codable consonants (i.e., /m/, /n/, /b/ or /p/, /d/ or /t/, /g/ or /k/, /w/, /l/, “y,” /s/, and “sh”) and/or a canonical syllable. Because the coding for the larger project focused only on communication, a fourth pass was used to code vocalizations that occurred outside of communication acts. Thus, the coder listened to the entire recording stopping it each time she heard a vocalization. For any vocalizations not already coded as part of a communication act, the coder marked the vocalization as a non-communicative vocalization and indicated whether it

included one or more codable consonants and/or a canonical syllable. SALT software was used to calculate the CSP and ECI vocal communication and vocal complexity variables. See Table 2 for the study variables.

For communicative vocalizations, two variables were calculated: the number of communication acts that include a vocalization and the proportion of vocalizations that are communicative (i.e., number of vocalizations within a communication act divided by the total number of vocalizations). For complexity, two additional variables were calculated: DKCC (Wetherby et al. 2007; Woynaroski et al. 2017) and the proportion of vocalizations with a canonical syllable (regardless of communicativeness).

Interobserver Reliability

A trained secondary coder independently coded a random sample of $\geq 20\%$ of coded sessions for each time point for CSP and ECI variables. The primary coder was blind to which sessions would be coded for reliability. Training included reading the coding manual and an initial training session with an expert coder including a didactic presentation, a question and answer session, and group coding of non-participants with discrepancy discussions. After the

Table 3 Operational definitions of key concepts for coding

Concept	Operational definition
Canonical syllable	Must include each of the following: (a) At least one consonant sound (i.e., /m/, /g/, /w/, “y,” “j,” “ng,” /v/, /n/, /p/, /t/, /k/, /l/, /s/, /t/, /z/, /f/, /r/, /b/, /d/, “zh,” “ch,” “sh,” and “th”) (b) At least one full vowel (c) Quick, uninterrupted transition from consonant to vowel or from vowel to consonant
Communication act	Behavior or set of behaviors must meet criteria for one of the following: (a) Word(s) (spoken or signed) (b) Nonword vocalization(s) with evidence of coordinated attention (c) One of the 15 specific gestures (i.e., tapping with fingers/hand, clapping, reaching, proximal pointing, distal pointing, “shh” gesture, head nod or head shake, wave, shoulder shrug, pantomime-like actions and depictive gestures, moving object toward adult, upturned palm, giving object, showing object, and hand as tool) with evidence of coordinated attention to message/referent and communication partner
Coordinated attention	Participant displays evidence of sequential or simultaneous attention to a person and an object or event within 3 s of his or her vocalization or gesture
Spoken word	Spoken words must meet the following criteria: (a) Represent a referent that is plausible within the context of the communication sample (b) Sufficiently approximate the adult pronunciation of the word
Vocalization	Nonvegetative voiced sound (i.e., one that is created by vibrating vocal folds) created during exhalation (eggressive phonation) Spoken words are coded as vocalizations and also included as spoken words to allow separate analyses The following sounds are not coded as vocalizations: <ul style="list-style-type: none"> • Voiced laughs, voiced sighs, and voiced cries because they are difficult to differentiate from other non-communicative noises • Whispered productions because they do not include voicing • Isolated voiceless consonants (e.g., /f/, /s/, /k/, /t/, /p/, and “sh”) because they do not include voicing • Glottal fry because the false vocal folds, not true vocal folds, are used • Ingressive phonation (i.e., vocalizations made when inhaling) • Reflexive, vegetative sounds resulting from burps, hiccups, coughs, sneezes, throat clearings, clicking of tongue, and popping of lips

initial training session, coders independently coded novel videos and participated in discrepancy discussions until the secondary coder reached criterion of at least .80 small/large agreement for all variables on three consecutive videos (Yoder et al. 2018). After initial training was complete, coders completed discrepancy discussions for each reliability set (i.e., group of five videos from which one reliability video was randomly chosen and completed for reliability before proceeding to the next set) to prevent coder drift. The primary coder's coding was used in the analyses. Interobserver reliability was estimated using intraclass correlation coefficients (ICCs) with absolute agreement and participant and observer as random factors. ICCs account for differences in unitizing and classifying behaviors between coders and for the variance among participants on the component variables addressing the research questions.

Results

Preliminary Analyses

Reliability

For all conventionally-coded variables combined, the mean ICC was .93 ($SD = .11$). Table 4 displays ICCs for the conventionally-coded variables by time period and procedure. Means and standard deviations are reported for ECI ICCs because these values were calculated from months 1 through 3 for Time 1 and months 11 through 13 for Time 3. We used a benchmark of .70 when interpreting the ICCs, which Mitchell (1979) interpreted as “very good”.

Creating Composite Variables

Six composite variables were computed: one for expressive language, one for nonverbal cognitive skills, and four for the vocal variables if the intercorrelation among component variables posited to measure the same construct warranted it. To allow composite scores to show growth, we calculated

and averaged the z -scores for each component variable using the sample's Time 3 mean and standard deviation.

The expressive language component variables (see Table 2) correlated with each other at $r \geq .40$, our a priori criterion for aggregating, at both time periods. The nonverbal cognitive skills component variables correlated with each other at $r \geq .40$ at each time point, except for Time 2 VABS fine motor skills and Time 2 MSEL visual recognition subscale ($r = .37$). Because these components correlated sufficiently at Times 1 and 3 and with all other component variables at all time points, both were retained for the Time 2 composite. For the four vocal variables, we created composite variables across the CSP and ECI sampling procedures for Time 1 and another composite for Time 3. For each vocal variable, we summed the results from the ECI Month 1, ECI Month 2, and ECI Month 3 for the Time 1 ECI value. We then averaged the Time 1 ECI value with the Time 1 CSP. Analogously, we summed the results from the ECI Months 11, 12, and 13 for the Time 3 ECI value. We then averaged the Time 3 ECI value with the Time 3 CSP. Correlations between Time 1 CSP and Time 1 ECI as well as Time 3 CSP and Time 3 ECI correlated at $r \geq .60$ for all four vocal variables, which is above the $r \geq .40$ criterion.

Evaluating Evidence of Convergent Validity

For evidence of convergent validity we tested whether each vocal variable predicted later expressive language skills using growth curve modeling with full maximum likelihood estimation (Enders 2010). By centering time in study at Time 3, the intercepts of the growth model are interpretable for the participants' expressive language skills at the final study period. Significant fixed coefficients for the predictor variable of a model with the expressive language composite as the dependent variable provide evidence of convergent validity.

The initial step of multilevel modeling is to identify the unconditional growth model. We used a build-up approach for model selection. Although the data fit the random intercept, random slope model better than the random intercept,

Table 4 Intraclass correlation coefficients for conventionally-coded vocal variables by time and procedure

Variable	Time 1		Time 3	
	CSP	ECI Mean (SD)	CSP	ECI Mean (SD)
Number of communication acts that include a vocalization	.97	.97 (.01)	.99	.98 (.02)
Proportion of vocalizations that are communicative	.84	.84 (.17)	.98	.96 (.03)
Diversity of key consonants used in communication acts	.79	.90 (.02)	.98	.96 (.02)
Proportion of vocalizations with a canonical syllable	.96	.92 (.06)	.97	.92 (.04)

CSP Communication Sample Procedure; ECI Early Communication Index (Greenwood et al. 2006; Luze et al. 2001); Time 1 study initiation; Time 3 12 months after study initiation; ECI intraclass correlation coefficients are reported as means of months 1–3 for Time 1 and months 11–13 for Time 3

fixed slope model ($p < .001$), the correlation between slope and intercept was very high ($r = .92$). Due to this high covariance of the intercept and slope and the desire to use the most parsimonious growth model, we chose to use the random intercept, fixed slope model. The growth parameter of interest was intercept, which was interpreted as the best estimate of end point (Time 3) expressive language.

To evaluate for evidence of convergent validity, we added each vocal variable to the random intercept, fixed slope model predicting the end point-centered intercept of growth of expressive language (i.e., best estimate of Time 3 expressive language) one at a time. All variables were significant predictors with positive associations (see Table 5). See Supplemental Material Fig. 1 for an example depiction of the association between a vocal variable and end point expressive language. Specifically, the association between DKCC at Time 3 and the predicted Time 3 expressive language value is displayed. No evidence of heteroscedasticity was observed. All residuals fell within the acceptable parameters for skewness (< 1.81) and kurtosis (< 13.01 ; Tabachnick and Fidell 2001). Mathematically pseudo R^2 is the difference between the residual variance of the intercept between the full model (i.e., includes vocal predictor variable of interest) and the reduced model (i.e., excludes vocal predictor variable of interest) divided by the reduced model’s residual variance of the intercept. Conceptually, pseudo R^2 means the proportion of the growth model or growth parameter explained by the more elaborate or full model relative to the less elaborate or unconditional model. Using pseudo $R^2 \geq .25$ as an indication of a large effect size, all four vocal variables at Time 1 have a large association with the best estimate of later expressive language.

Evaluating Evidence of Divergent Validity

For evidence of divergent validity, we evaluated whether each vocal development variable predicted the end point-centered intercept of the growth of nonverbal cognitive skills (i.e., best estimate of Time 3 nonverbal cognitive skills). Theoretically, vocalization measures should *not* be significant predictors of nonverbal cognitive skills. As before, time in study is centered at Time 3 to yield intercepts interpretable for the participants’ skills at the final study period.

Associations with the intercept of growth on nonverbal cognitive skills are expected to be low and nonsignificant. Given the large sample size, we rely on significance to define low.

We used a build-up approach for model selection for the model predicting nonverbal cognitive skills. As with expressive language, the data fit the random intercept, random slope model better than the random intercept, fixed slope model ($p < .001$). Unlike the expressive language model, the intercorrelation between the intercept and slope for the random intercept, random slope model was acceptably low ($r = .79$) to retain the random slope. To reflect the logic used in the expressive language models, the intercept was the parameter of interest. None of the vocal variables were significant predictors of end point nonverbal cognitive skills (see Table 6). No evidence of heteroscedasticity was observed. All residuals fell within the acceptable parameters for skewness and kurtosis (Tabachnick and Fidell 2001).

Evaluating Evidence of Sensitivity to Change

A significant difference between Time 1 and Time 3 via a paired *t*-test is evidence of sensitivity to change (see Table 7). All of the variables exhibited evidence of sensitivity to change and exceeded the benchmark for large effect size (i.e., Cohen’s $d \geq .80$).

Discussion

Summary of Relative Validity of Vocal Variables

The two vocal variables purported to assess communicative vocalizations (i.e., number of communication acts that include a vocalization and proportion of vocalizations that are communicative) and the two purported to assess vocal complexity (i.e., DKCC and proportion of vocalizations with a canonical syllable) exhibited consistent evidence for convergent validity, divergent validity, and sensitivity to change. Regardless of whether we use the presence or absence of significant results or effect size as the criterion for assigning validity and summarizing across these three purposes, all four variables presented with consistent positive evidence for all three purposes. As further indication of strong validity

Table 5 Fixed effects estimates for vocal variables predicting end point expressive language

Vocal Variable	Coeff.	SE	<i>t</i>	<i>df</i>	<i>p</i>	Pseudo R^2
Number of communication acts that include a vocalization	0.86	0.10	8.64	86.16	<.001	.54
Proportion of vocalizations that are communicative	0.69	0.07	9.35	85.55	<.001	.59
Diversity of key consonants used in communication acts	0.59	0.06	9.41	85.86	<.001	.60
Proportion of vocalizations with a canonical syllable	0.38	0.05	7.26	86.20	<.001	.44

Coeff coefficient value; Pseudo $R^2 = 1 - (\text{covariance parameter for intercept of model with the vocal variable of interest} / \text{covariance parameter for intercept of model without the vocal variable of interest})$

Table 6 Fixed effects estimates for main effects of vocal variables predicting end point nonverbal cognitive skills

Vocal variable	Coeff.	SE	<i>t</i>	<i>df</i>	<i>p</i>
Number of communication acts that include a vocalization	0.06	0.11	0.54	86.09	.59
Proportion of vocalizations that are communicative	−0.06	0.09	−0.75	86.85	.46
Diversity of key consonants used in communication acts	0.02	0.07	0.25	86.14	.81
Proportion of vocalizations with a canonical syllable	−0.04	0.05	−0.70	86.14	.49

Coeff coefficient value

Table 7 Results of paired *t*-tests from Time 1 to Time 3

Vocal variable	Mean	<i>SD</i>	95% CI	<i>t</i>	<i>d</i>
Number of communication acts that include a vocalization	0.73	0.75	[0.56, 0.89]	8.70	0.88***
Proportion of vocalizations that are communicative	0.79	0.77	[0.61, 0.96]	9.16	0.91***
Diversity of key consonants used in communication acts	1.13	0.79	[0.95, 1.30]	12.80	1.32***
Proportion of vocalizations with a canonical syllable	1.00	1.03	[0.77, 1.22]	8.72	1.03***

*** = $p < .001$; *p* values are for two-tailed significance tests; *d* = within subjects effect size accounting for correlation between Time 1 and Time 3

evidence, the effect sizes ranged from pseudo $R^2 = .44$ to $.60$ for convergent validity and $d = 0.88$ to 1.32 for sensitivity to change. DKCC had the largest effect size for convergent validity and sensitivity to change. Large effect sizes increase our confidence that the findings will replicate. The convergence across validity comparison methods bolsters the confidence in the conclusions.

The sensitivity to change evidence provides key information for planning intervention studies. For example, using a vocal variable to show response to intervention or as a mediator to explain why a given intervention is effective in facilitating language in preverbal children with ASD, one must select a vocal variable that captures change over time.

The Current Study Findings Relative to the Extant Literature

The social feedback theory, social feedback loop, transactional theory of spoken language development, and child-driven theories all support using vocal communication and vocal complexity variables as putative predictors of expressive language. Having multiple studies suggesting that a variable or construct should predict language strengthens the rationale for using that variable or construct. The current study was not designed to test one theoretical framework against another. All of these theories predict that vocal communication and vocal complexity should change over time. Additionally, the bidirectional theories suggest that increases in child vocal communication and child vocal complexity will elicit more frequent or complex adult responses to scaffold the child's ability to produce more adult-like productions including spoken words. Future work will be needed to test this prediction. The construct validity evidence in which vocal communication and vocal complexity variable

predict expressive language but not later nonverbal cognition supports the assertions that the examined measures of vocal communication and vocal complexity measure the constructs they are presumed to measure.

Convergent validity evidence based on significant correlations with current or later expressive language has been reported for vocal communication (Plumb 2008) and DKCC (Wetherby et al. 2007; Woynaroski 2014; Woynaroski et al. 2017; Yoder et al. 2015a). Thus, these findings from the current study are replications, decreasing the likelihood that the current study findings are sample specific results.

Some of the many tests of significance for the association between vocal communication and language are significant, while others are nonsignificant. For example, Plumb and Wetherby (2013) found that the proportion of vocalizations that are communicative was associated with the Speech subscale of the Communication and Symbolic Behavior Scales, a measure of speech level, but unrelated to the MSEL verbal developmental quotient, a measure of language delay. Swineford (2011) found that rate of vocal communication acts was unrelated to expressive language. The current study may have identified these previously unidentified relations in part due to different metrics for the vocal variable, different metrics for the language variable, the relatively large sample size in the current study, use of growth curve modeling to generate a better estimate of end point expressive language rather than a single observed measure (Singer and Willett 2003), and the focus on expressive language level rather than expressive language delay. In comparison to the current study, Plumb and Wetherby (2013) and Swineford (2011) included fewer participants and quantified expressive language with a single measure rather than multiple measures.

No known prior studies report convergent validity for the proportion of vocalizations that include a canonical syllable

in children with ASD. However, Williams (2013) reported a nonsignificant relation between a similar variable (i.e., percent of syllables that are canonical) and language composites for fifteen 6-month-old siblings of children with ASD. The fact that most infant siblings will not have ASD, the relatively younger and smaller sample, and the use of different language measures relative to the current study may explain the incongruent findings.

No prior studies reporting evidence for divergent validity for any of the study variables with children with ASD were located. For sensitivity to change, only one study with children with ASD was located. Woynaroski et al. (2016) reported significant positive simple linear growth in DKCC for 87 initially preverbal children with ASD (mean age = 34.7 months, SD = 7.2 months) across 16 months in a longitudinal correlational study. Thus, the evidence for sensitivity to change for DKCC in the current sample replicates this finding in a more diverse sample of participants with ASD.

Limitations

Four limitations should be acknowledged. First, validation refers to a specific variable, use, and population (Yoder et al. 2018). Therefore, findings from this study may not directly transfer to other variables derived from the same data collection methods, other uses, or other populations. Second, multiple *t*-tests were conducted without alpha adjustment when assessing significance of predicted associations and change, which increases the risk for Type I errors. Although using composites partially addresses concerns about family-wise error due to multiple significance tests without alpha adjustment, there are still many significance tests per research question. Replications of associations with expressive language that are new to the field are needed to ensure those findings are not sample specific. Despite some novel findings, many of the predictors of expressive language have been detected in other samples of children with ASD, as described above. It is unlikely that replicated associations are significant due to unadjusted multiple significance testing. Third, clear divisions among variables that measure vocal communication and vocal complexity were not possible in every case. For example, DKCC considers vocal communication and complexity. Last, the correlational and single-group pre-post design used to test the validity of the selected vocal variables prevents confident inferences that predictors cause criterion variables or that treatments caused the change in the vocal variables.

Strengths

Five strengths should be acknowledged. First, we addressed not only convergent validity, but also divergent validity when

assessing construct validity. Divergent validity evidence is notably sparse in the literature. Thus, the current findings provide a unique contribution to the literature, particularly for vocal measures for children with ASD. Second, we used multilevel modeling to provide the best estimate of end point expressive language and end point nonverbal cognitive skills, rather than relying on the observed value (Singer and Willett 2003). Third, the study duration of 12 months provides a relatively long, and meaningful, period of time for predicting growth. Intervention goals are often written for yearlong intervals. The current study design permitted addressing predictive validity, one important purpose for which vocal variables are often needed. Fourth, this study includes a relatively large sample size for this population, which increased the power to detect effects and permitted the use of multilevel models with the necessary number of predictors to address the research questions. Fifth, composite variables were used to increase the reliability of construct estimates relative to single-measure constructs.

Implications

The results provide guidance for selecting variables for a variety of studies related to vocal development and language development of children with ASD. Overall, the results support the measurement of vocal communication and vocal complexity when derived by human coding of communication samples to assess vocal level and vocal development in young children with ASD. Potential purposes for assessing and targeting vocal development include increasing the effectiveness of communication intervention, identifying early response to such interventions, and explaining why or for whom communication interventions are effective in preschool-aged children with ASD. For example, when selecting variables that might mediate treatment effects on expressive language, the findings suggest that using vocal communication and vocal complexity variables may maximize the probability of detecting the putative mediated effect of early language interaction on expressive language through midpoint vocal development.

Future Directions

Additional investigations are needed to compare the validity of these vocal variables to less costly variables, including automated measures. Using variables that are more elaborate and/or expensive due to research staff training and coding costs is justifiable only when the more elaborate or more costly vocal variables yield more useful results than the less elaborate or less costly measures. Assessing volubility (i.e., how frequently a child vocalizes) is less elaborate than vocal communication and vocal complexity variables, and therefore less expensive to code

(i.e., lower training and coding costs). Automated measures are also less expensive to code than variables coded conventionally from communication samples. Examples of automated variables from day-long audio samples from the Language ENvironment Analysis (LENA) system for volubility or complexity include the number of child speech-related vocalizations (LENA 2015), the average count per utterance of consonants and vowels (Woynaroski et al. 2017; Xu et al. 2014), and the infraphonological vocal development (IVD) score (Oller et al. 2010). Whether conventionally-coded variables account for unique variance above and beyond less elaborate or less expensive measures should be evaluated empirically.

In addition, further investigation is required to validate the tested variables with other populations of children, such as children with language impairment without ASD or children with ASD at different communication levels. Relatedly, because construct validity is judged based on a network of constructs called nodes (Cronbach and Meehl 1955), additional nodes for convergent and divergent validity may be explored to increase the confidence in the results. For instance, correlations with receptive language may be considered for convergent validity nodes. It may not be intuitive to hypothesize that vocal variables may predict receptive language, but receptive language has predicted DKCC (Woynaroski et al. 2016). One possible explanation for this finding is that children may be trying to use words they understand prior to being able to make themselves understood. Other divergent validity nodes could also be considered.

Whether variables that may most readily transfer to clinical practice (e.g., DKCC and proportion of vocalizations with canonical syllables) can be coded live reliably, and with what amount of training, warrants further investigation. If these variables can be coded reliably, their use may be encouraged within clinical and research settings with appropriate training.

Conclusion

The current study offers crucial new knowledge that can help the broader scientific community measure vocal development within and across young children with ASD. Key findings include strong evidence of convergent validity, divergent validity, and sensitivity to change for predicting expressive language using conventional methods for coding communication samples to measure vocal communication and vocal complexity in young children with ASD. These results support the use of conventional measures of vocal communication and vocal complexity in future studies of communication intervention in children with ASD.

Author Contributions JM conceived of the study, participated in the design, coded the data, interpreted the data, and drafted the manuscript; PY participated in the design, helped interpret the data and draft the manuscript, and was a site PI on the grant. AE and SR were a site and overall PI, respectively, on the grant, participated in the overall design of the overarching study, led recruitment efforts, supported interventions for the participants and their families over the course of the study, and edited the manuscript. All authors read and approved the final manuscript.

Funding This research was funded by the NIMH (5R01MH100030; PI: Rogers) and supported by a US Department of Education Preparation of Leadership Personnel grant (H325D140087) and the EKS of NIH (U54HD083211). We thank all of the children and families who participated to make this work possible. We also thank Kadie Ulven for coding diligently for this project.

Compliance with Ethical Standards

Conflicts of interest SR has potential conflicts of interest related to publications related to interventions used in this study for which she receives royalties. However, all royalties from those publications were donated to autism research. JM, PY, and AE declare that they have no conflict of interest.

Ethical Approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed Consent Caregivers provided written consent for participants included in the study.

References

- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.
- Billstedt, E., Gillberg, C., & Gillberg, C. (2005). Autism after adolescence: Population-based 13- to 22-year follow-up study of 120 individuals with autism diagnosed in childhood. *Journal of Autism and Developmental Disorders*, *35*, 351–360. <https://doi.org/10.1007/s10803-005-3302-5>.
- Book, L. A. (2009). *Early red flags for autism spectrum disorders in toddlers in the home environment* (Doctoral dissertation). Retrieved from ProQuest database. (UMI No. 3399180).
- Camarata, S., & Yoder, P. (2002). Language transactions during development and intervention: Theoretical implications for developmental neuroscience. *International Journal of Developmental Neuroscience*, *20*, 459–465. [https://doi.org/10.1016/S0736-5748\(02\)00044-8](https://doi.org/10.1016/S0736-5748(02)00044-8).
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*, 81–105. <https://doi.org/10.1037/h0046016>.
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*, 281–302.
- Enders, C. K. (2010). *Applied missing data analysis*. New York: The Guilford Press.
- Fagan, M. K., & Doveikis, K. N. (2017). Ordinary interactions challenge proposals that maternal verbal responses shape infant vocal development. *Journal of Speech, Language, and Hearing*

- Research, 60, 2819–2827. https://doi.org/10.1044/2017_JSLHR-S-16-0005.
- 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: Brookes.
- Fry, D. (1966). The development of the phonological system in the normal and deaf child. In F. Smith & G. A. Miller (Eds.), *The genesis of language* (pp. 187–206). Cambridge, MA: MIT Press.
- Goldstein, M. H., King, A. P., & West, M. J. (2003). Social interaction shapes babbling: Testing parallels between birdsong and speech. *Proceedings of the National Academy of Sciences United States of America*, 100, 8030–8035. <https://doi.org/10.1073/pnas.1332441100>.
- Goldstein, M. H., & Schwade, J. A. (2008). Social feedback to infants' babbling facilitates rapid phonological learning. *Psychological Science*, 19, 515–523. <https://doi.org/10.1111/j.1467-9280.2008.02117.x>.
- Greenwood, C. R., Carta, J. J., Walker, D., Hughes, K., & Weathers, M. (2006). Preliminary investigations of the application of the Early Communication Indicator (ECI) for infants and toddlers. *Journal of Early Intervention*, 28, 178–196. <https://doi.org/10.1177/105381510602800306>.
- Greenwood, C. R., Walker, D., & Buzhardt, J. (2010). The early communication indicator for infants and toddlers: Early Head Start growth norms from two states. *Journal of Early Intervention*, 32, 310–334. <https://doi.org/10.1177/1053815110392335>.
- Howlin, P. (2000). Outcome in adult life for more able individuals with autism or Asperger syndrome. *Autism*, 4, 63–83. <https://doi.org/10.1177/1362361300004001005>.
- Iverson, J. M. (2010). Developing language in a developing body: The relationship between motor development and language development. *Journal of Child Language*, 37, 229–261. <https://doi.org/10.1017/S0305000909990432>.
- LENA Research Foundation. (2015). LENA Pro. Retrieved from <http://www.lenafoundation.org/lena-pro>.
- Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism Diagnostic Interview-Revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, 24, 659–685. <https://doi.org/10.1007/bf02172145>.
- Luyster, R., Gotham, K., Guthrie, W., Coffing, M., Petrak, R., Pierce, K., ... Lord, C. (2009). The Autism Diagnostic Observation Schedule—Toddler Module: A new module of a standardized diagnostic measure for autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 39, 1305–1320. <https://doi.org/10.1007/s10803-009-0746-z>.
- Luze, G. J., Linebarger, D. L., Greenwood, C. R., & Carta, J. J. (2001). Developing a general outcome measure of growth in the expressive communication of infants and toddlers. *School Psychology Review*, 30, 383–406.
- McCoy, D. (2013). *Observation of social communication red flags in young children with autism spectrum disorder, developmental delay, and typical development using two observation methods* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses Global database. (UMI No. 3596541 Ph.D).
- McDaniel, J., D'Ambrose Slaboch, K., & Yoder, P. (2018). A meta-analysis of the association between vocalizations and expressive language in young children with autism spectrum disorder. *Research in Developmental Disabilities*, 72, 202–213. <https://doi.org/10.1016/j.ridd.2017.11.010>.
- 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. [https://doi.org/10.1044/1092-4388\(2009/09-0023\)](https://doi.org/10.1044/1092-4388(2009/09-0023)).
- McLean, J., & Snyder-McLean, L. (1978). *A transactional approach to early language training*. Columbus, OH: Charles E. Merrill.
- Miller, J., & Chapman, R. (2016). *Systematic Analysis of Language Transcripts (Version 16) [Computer software]*. Madison, WI: SALT Software LLC.
- Mitchell, S. K. (1979). Interobserver agreement, reliability, and generalizability of data collected in observational studies. *Psychological Bulletin*, 86, 376–390. <https://doi.org/10.1037/0033-2909.86.2.376>.
- Mullen, E. M. (1995). *Mullen Scales of Early Learning* (AGS ed.). Los Angeles: Western Psychological Services.
- Oller, D. K. (2000). *The emergence of the speech capacity*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Oller, D. K., Eilers, R. E., Neal, A. R., & Schwartz, H. K. (1999). Precursors to speech in infancy: The prediction of speech and language disorders. *Journal of Communication Disorders*, 32, 223–245. [https://doi.org/10.1016/S0021-9924\(99\)00013-1](https://doi.org/10.1016/S0021-9924(99)00013-1).
- Oller, D., Niyogi, P., Gray, S., Richards, J., Gilkerson, J., Xu, D., ... Warren, S. (2010). Automated vocal analysis of naturalistic recordings from children with autism, language delay, and typical development. *Proceedings of the National Academy of Sciences United States of America*, 107, 13354–13359. <https://doi.org/10.1073/pnas.1003882107>.
- Plumb, A. M. (2008). *Vocalizations of children with autism spectrum disorders late in the second year of life* (Doctoral dissertation). Retrieved from ProQuest database. (UMI No. 3348528).
- Plumb, A. M., & Wetherby, A. M. (2013). Vocalization development in toddlers with autism spectrum disorder. *Journal of Speech, Language, and Hearing Research*, 56, 721–734. [https://doi.org/10.1044/1092-4388\(2012/111-004\)](https://doi.org/10.1044/1092-4388(2012/111-004)).
- Rogers, S. J., Estes, A., & Yoder, P. (2013). Intervention effects of intensity and delivery style for toddlers with ASD (5R01MH100030), National Institute of Mental Health.
- Sameroff, A., & Chandler, M. (1975). Reproductive risk and the continuum of caretaking casualty. In M. Horowitz, M. Hetherington, S. Scarr-Salapatek, & G. Siegel (Eds.), *Review of child development research* (pp. 187–244). Chicago: University Park Press.
- Sheinkopf, S. J., Mundy, P., Oller, D. K., & Steffens, M. (2000). Vocal atypicalities of preverbal autistic children. *Journal of Autism and Developmental Disorders*, 30, 345–354. <https://doi.org/10.1023/A:1005531501155>.
- Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York: Oxford University Press.
- Sparrow, S. S., Cicchetti, D. V., & Balla, D. A. (2005). *Vineland Adaptive Behavior Scales* (2nd ed.). Circle Pines, MN: American Guidance Service.
- Stark, R. E., & Tallal, P. (1981). Selection of children with specific language deficits. *Journal of Speech and Hearing Disorders*, 46, 114–122. <https://doi.org/10.1044/jshd.4602.114>.
- Stoel-Gammon, C. (2011). Relationships between lexical and phonological development in young children*. *Journal of Child Language*, 38, 1–34. <https://doi.org/10.1017/S0305000910000425>.
- Swineford, L. B. (2011). *Symbol use in the home environment in toddlers suspected of having autism spectrum disorder* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses Global database. (UMI No. 3502964).
- Tabachnick, B., & Fidell, L. (2001). *Using multivariate statistics* (4th ed.). Boston: Allyn and Bacon.
- Talbott, M. R. (2014). *Autism risk status and maternal behavior: Impacts on infant language and communication development from 6 to 36 months of age* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses Global database. (UMI No. 3625962).
- Tapp, J. (2003). *ProCoderDV*. Nashville, TN: Vanderbilt Kennedy Center.

- Vihman, M. (1992). Early syllables and the construction of phonology. In C. A. Ferguson, L. Menn, & C. Stoel-Gammon (Eds.), *Phonological development: Models, research, implications* (pp. 393–422). Timonium, MD: York Press.
- Vihman, M. (1996). *Phonological development: The origins of language in the child*. Cambridge, MA: Blackwell Publishing Ltd.
- 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. <https://doi.org/10.1177/0956797614531023>.
- West, M. J., & Rheingold, H. L. (1978). Infant stimulation of maternal instruction. *Infant Behavior and Development*, 1, 205–215. [https://doi.org/10.1016/S0163-6383\(78\)80031-9](https://doi.org/10.1016/S0163-6383(78)80031-9).
- Wetherby, A. M., Watt, N., Morgan, L., & Shumway, S. (2007). Social communication profiles of children with autism spectrum disorders late in the second year of life. *Journal of Autism and Developmental Disorders*, 37, 960–975. <https://doi.org/10.1007/s10803-006-0237-4>.
- Wetherby, A. M., Yonclas, D. G., & Bryan, A. A. (1989). Communicative profiles of preschool children with handicaps: Implications for early identification. *Journal of Speech and Hearing Disorders*, 54, 148–158. <https://doi.org/10.1044/jshd.5402.148>.
- Williams, L. (2013). *Vocalization behavior during the autism observation scale for infants in siblings of children with autism spectrum disorders* (unpublished Master's thesis). University of Washington: Seattle, WA.
- Woynaroski, T. (2014). *The stability and validity of automated vocal analysis in preschoolers with autism spectrum disorder in the early stages of language development* (Doctoral dissertation). Retrieved from ProQuest database. (UMI No. 3648771).
- Woynaroski, T., Oller, D. K., Keceli-Kaysili, B., Xu, D., Richards, J. A., Gilkerson, J., ... Yoder, P. (2017). The stability and validity of automated vocal analysis in preverbal preschoolers with autism spectrum disorder. *Autism Research*, 10, 508–519. <https://doi.org/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. <https://doi.org/10.1007/s10803-015-2647-7>.
- Woynaroski, T., Yoder, P. J., Fey, M. E., & Warren, S. F. (2014). A transactional model of spoken vocabulary variation in toddlers with intellectual disabilities. *Journal of Speech, Language, and Hearing Research*, 57, 1754–1763. https://doi.org/10.1044/2014_JSLHR-L-13-0252.
- Xu, D., Richards, J. A., & Gilkerson, J. (2014). Automated analysis of child phonetic production using naturalistic recordings. *Journal of Speech, Language, and Hearing Research*, 57, 1638–1650. https://doi.org/10.1044/2014_JSLHR-S-13-0037.
- Yoder, P. J., & Feagans, L. (1988). Mothers' attributions of communication to prelinguistic behavior of developmentally delayed and mentally retarded infants. *American Journal on Mental Retardation*, 93, 36–43.
- Yoder, P. J., Lloyd, B. P., & Symons, F. J. (2018). *Observational measurement of behavior* (2nd ed.). Baltimore: Paul H. Brooks Publishing Company.
- Yoder, P. J., & Warren, S. F. (2002). Effects of prelinguistic milieu teaching and parent responsivity education on dyads involving children with intellectual disabilities. *Journal of Speech, Language, and Hearing Research*, 45, 1158–1174. [https://doi.org/10.1044/1092-4388\(2002\)094](https://doi.org/10.1044/1092-4388(2002)094).
- Yoder, P., Watson, L., & Lambert, W. (2015a). 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. <https://doi.org/10.1007/s10803-014-2286-4>.
- Yoder, P. J., Woynaroski, T., Fey, M. E., Warren, S. F., & Gardner, E. (2015b). Why dose frequency affects spoken vocabulary in preschoolers with Down syndrome. *American Journal on Intellectual and Developmental Disabilities*, 120, 302–314. <https://doi.org/10.1352/1944-7558-120.4.302>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.