Evaluating the Validity of Vocalization Measures for Assessing Vocal Development in Young Children with Autism Spectrum Disorder

By

Jena McDaniel

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

In partial fulfillment of the requirements

For the degree of

DOCTOR OF PHILOSOPHY

in

Hearing and Speech Sciences

May 10, 2019

Nashville, Tennessee

Approved:

C. Melanie Schuele, Ph.D.
Paul Yoder, Ph.D.
Tiffany Woynaroski, Ph.D.
Stephen Camarata, Ph.D.

Copyright © 2019 by Jena McDaniel
All Rights Reserved

To Mom, Dad, and my brother Jacob

ACKNOWLEDGEMENTS

This work was funded by the National Institute of Mental Health (5R01MH100030; PI: Rogers) and supported by a US Department of Education Preparation of Leadership Personnel grant (H325D140087; PI: Schuele). The content is solely the responsibility of the author(s) and does not necessarily represent the official views of the National Institute of Mental Health or the US Department of Education.

I greatly appreciate all of the children and families who participated in this study and made this work possible. I would like to thank all of my mentors, during my undergraduate, graduate, and clinical training, for patiently sharing their time and expertise with me to support my goals for serving children and families. I also thank my colleagues who have participated in this work and related projects throughout my education and professional career for pushing me to address clinically relevant questions.

I am deeply thankful for the children and families I have met and been fortunate to work with as a student and as a speech-language pathologist for providing continuous motivation to discover how to enable children to reach their potential with enthusiasm and joy. Finally, thank you to my family and friends for caring for me well and for reminding me of my purpose and motivation over the years.

TABLE OF CONTENTS

Page

DEDICATION ... iii

ACKNOWLEDGEMENTS ... iv

LIST OF TABLES ... vii

LIST OF FIGURES ... viii

LIST OF ABBREVIATIONS ... ix

Chapter

I. Introduction ... 1

Vocal Development in Children with Typical Development ... 2
Potentially Important Aspects of Vocalizations for Children with ASD ... 4
Theoretical Support for Measuring Selected Aspects of Vocalizations ... 6
Social feedback theory ... 6
Speech attunement framework ... 7
Transactional theory of spoken language development ... 7
Application to vocal development in children with ASD ... 8
Empirical Support for Measuring Selected Aspects of Vocalizations ... 9
Volubility ... 10
Communicative use ... 11
Complexity ... 11
Vocal reciprocity ... 13
The Need for Additional Evidence of Validity for Vocal Variables ... 14
Comparing Validity Evidence Across Vocal Variables ... 15
Purpose and Research Questions ... 16

II. Method ... 18

Participants ... 18
Procedures ... 20
Day-long naturalistic audio samples ... 21
Communication Sample Procedure ... 22
Early Communication Index ... 22
MacArthur-Bates Communicative Development Inventory ... 22
Mullen Scales of Early Learning ... 23
Vineland Adaptive Behavior Scales, Second Edition ... 23
Observational Coding ... 23
Variables Derived from Observational Coding ... 26
Volubility variables ... 26
Communicative use of vocalizations variables ... 26
Vocal complexity variables ... 27
Automated Vocal Analysis ... 27
Variables Derived from Automated Vocal Analysis ... 29

Volubility variable 29
Vocal complexity variables........ 29
Vocal reciprocity variable 30
Interobserver Reliability 31

III. Results 33

Preliminary Analyses 33
Reliability........ 33
Eliminating redundant vocal variables 34
Creating composite variables........ 35
Evaluating Evidence of Convergent Validity 37
Evaluating Evidence of Divergent Validity 39
Evaluating Evidence of Sensitivity to Change........ 41
Evaluating Incremental Validity Relative to Volubility 41
Evaluating Incremental Validity of RVC Relative to Communicative Use and Complexity ... 44
Evaluating Incremental Validity of Conventionally-Coded Measures Relative to Automated Measures of the Same Construct 45

IV. Discussion........ 47

Summary of Relative Validity of Vocal Variables 47
The Current Study Findings Relative to the Extant Literature........ 51
Limitations........ 55
Strengths........ 56
Implications 57
Future Directions........ 58
Conclusion 59

Appendix

A. Intercorrelations Between Vocal Variables of Communicative Use and Complexity 60

B. Growth Curve Model Results for Convergent Validity Analyses 62

C. Growth Curve Model Results for Divergent Validity Analyses 66

D. Growth Curve Model Results for Incremental Validity Analyses........ 71

REFERENCES 78

LIST OF TABLES

Table Page

1. Participant Characteristics at Study Entry 20

2. Study Constructs, Procedures, and Variables 21

3. Operational Definitions of Key Concepts for Coding........ 25

4. Intraclass Correlation Coefficients for Conventionally-Coded Vocal Variables by Time and Procedure 34

5. Fixed Effect Estimates for Vocal Variables Predicting End Point Expressive Language...... 39

6. Fixed Effect Estimates for Main Effects of Vocal Variables Predicting End Point Nonverbal Cognitive Skills........ 40

7. Results of Paired *t*-Tests from Time 1 to Time 3 41

8. Unstandardized Coefficients (Standard Errors) and Significance for More Elaborate Vocal Variables Predicting End Point Expressive Language After Controlling for Conventionally-Coded Volubility 42

9. Unstandardized Coefficients (Standard Errors) and Significance for More Elaborate Vocal Variables Predicting End Point Expressive Language After Controlling for an Automated Volubility Variable 43

10. Unstandardized Coefficients (Standard Errors) and Significance for RVC, a Dyadic Three-Event Variable, Predicting End Point Expressive Language After Controlling for Single Actor / Single Event Variables........ 45

11. Unstandardized Coefficients (Standard Errors) and Significance for the Conventionally-Coded Vocal Variable Predicting End Point Expressive Language After Controlling for the Automated Vocal Variable 46

12. Summary of Evidence of Convergent Validity, Divergent Validity, and Sensitivity to Change 47

13. Summary of Evidence of Incremental Validity for Predicting Later Expressive Language ... 49

14. Replicated Findings Identified in Extant Literature and Current Study 52

LIST OF FIGURES

Figure Page

1. Contingency Table for Calculating Reciprocal Vocal Contingency 31

LIST OF ABBREVIATIONS

ACPU: Average Count Per Utterance

ACPU-C: Average Count Per Utterance – Consonants

ACPU-C+V: Average Count Per Utterance – Consonants and Vowels

ACPU-V: Average Count Per Utterance – Vowels

ASD: Autism spectrum disorder

AV: Adult vocalization

CA: Communication act

CHNSP: Number of child speech-related vocalizations

Coeff.: Coefficient

Comm.: Communicative use

CS: Canonical syllable

CSBS: Communication and Symbolic Behavior Scales

CSP: Communication Sample Procedure

CU: Child utterances

CUC: Child utterance clusters

CV: Child vocalization

CVI: Child vocal island

DKCC: Diversity of key consonants used in communication acts

ECI: Early Communication Index

ICC: Intraclass correlation coefficients

IVD: Infraphonological vocal development

ITS: Integrated Time Segments

LENA: Language ENvironment Analysis

Max: Maximum

MB-CDI: MacArthur-Bates Communicative Development Inventory: Expressive vocabulary compilation form

Min: Minimum

MSEL: Mullen Scales of Early Learning

RVC: Reciprocal vocal contingency

SALT: Systematic Analysis of Language Transcripts

SORF: Systematic Observation of Red Flags for Autism Spectrum Disorders in Young Children

SORF-Home: Systematic Observation of Red Flags for Autism Spectrum Disorders in Young Children at Home

SCU: Speech-related child utterance

SVI: Speech-related vocal island

TADPOLE: Toddlers with Autism: Developing Opportunities for Learning

TV: Total vocalizations

VABS: Vineland Adaptive Behavior Scales

CHAPTER 1

INTRODUCTION

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 is important for children with ASD because language skills predict social, adaptive, and vocational outcomes in this population (Billstedt, Gillberg, & Gillberg, 2005; Howlin, 2000). For children with ASD, assessing and targeting vocal development, which is the process through which children produce increasingly speech-like sounds (Oller, 2000), may be useful for four 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 if children are not "ready" for linguistic targets. Second, vocal development might indicate early response to intervention that targets lexical development. Third, vocal development might help explain why language intervention is effective in facilitating language in initially preverbal children with ASD. Finally, pretreatment vocal development might describe for whom communication intervention is effective in initially preverbal children with ASD. However, to experimentally evaluate these potential purposes for focusing on vocal development, researchers must be able to measure vocal development in a valid manner.

Identifying valid vocal development measures for children with ASD is an essential prerequisite step to developing and evaluating interventions that may promote prelinguistic vocalizations and later language skills for children with ASD. Because there is no gold standard vocal development measure, one cannot 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 & Meehl, 1955). The strength of validity evidence, including construct validity, sensitivity to change, and incremental validity, influences the scientific usefulness of specific variables for specific purposes in research as well as clinical practice. This study assesses the validity of multiple variables purported to capture vocal development of children with ASD.

In this introduction, we briefly describe vocal development in children with typical development and the continuity of babbling with spoken language. Then, we introduce various aspects of vocalizations that warrant further investigation for children with ASD. Next, we present theoretical and empirical support for investigating the validity of variables that assess these aspects of vocalizations. We conclude by explaining the need for additional evidence to measure vocal development in a valid manner for children with ASD and how the current study addresses that need with a large, longitudinal sample of young children with ASD.

Vocal Development in Children with Typical Development

Vocal development is critical for spoken language (Oller, 2000). Here, we define vocalizations as nonvegetative voiced sounds (i.e., created by vibrating vocal folds) created during exhalation (i.e., eggressive phonation) because English phonemes are produced by eggressive phonation. Children with typical development produce a variety of vocalizations before using spoken words as well as when they are producing a relatively small number of spoken words. Initially vocalizations are reflexive and primitive. Over time they become intentional, more speech-like, and are used for communicative purposes (Oller, 2000; Oller, Eilers, Neal, & Schwartz, 1999). 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 canonical babbling includes vowel-like and consonant-like sounds with rapid, adult-like transitions between them

(Oller et al., 1999). In typical development first words emerge around 12 months of age. As children's expressive vocabulary sizes increases, they use a combination of babbling and speech during the second year of life. Also in the second year of life, vocal complexity increases, including consonant and syllable shape diversity as well as the ratio of words to nonwords (Wetherby, Cain, Yonclas, & Walker, 1988).

Current evidence supports the continuity of babbling and spoken words in typical development, in contrast to the discontinuity theories posited in the past (e.g., Jakobson, 1968). For example, for individual children, the phonemes produced in babbling appear in early words productions more so than phonemes that were not produced in the babbling (McCune & Vihman, 2001; Oller, 2000; Vihman, 2017; Vihman, Macken, Miller, Simmons, & Miller, 1985). After 10 months of age, the acoustic characteristics of babbling vary across spoken languages (Oller, 2000; Rvachew, Mattock, Polka, & Ménard, 2006). The language-specific nature of babbling suggests that babbling and spoken words are related rather than independent developmental processes. Additionally, how frequently children with typical development vocalize (i.e., volubility) and the complexity of their vocalizations (e.g., inclusion of consonants and canonical syllables) correlate with later expressive language measures (e.g., Stoel-Gammon, 1991; Watt, Wetherby, & Shumway, 2006). For example, consonant inventory close to age 2 (mean age = 20 months) predicts unique variance in expressive language scores on the Mullen Scales of Early Learning (MSEL; Mullen, 1995) above and beyond acts for joint attention and gesture inventory (Watt et al., 2006; Wetherby & Prizant, 2002). At 24 months of age, the number of consonants produced in the initial word position and the number of consonants produced in the final word position correlate with the concurrent number of different words produced, $r = .79$, $p < .001$, and $r = .85$, $p < .001$, respectively (Stoel-Gammon, 1989, 1991).

The influence of social interactions and reciprocity in vocal development has been investigated. The social feedback theory asserts that, "infants' prelinguistic vocalizations, and caregivers' reactions to those immature sounds, create opportunities for social learning that afford infants knowledge of phonology" (Goldstein & Schwade, 2008, p. 522). Caregivers respond contingently to infants' prelinguistic vocalizations based on the features and context of those vocalizations (Goldstein & West, 1999; Gros-Louis, West, Goldstein, & King, 2006) in ways that appear to support language development. For example, Gros-Louis et al. (2006) found that mothers imitated their children's consonant-vowel vocalizations more frequently than their vowel-like vocalizations. Goldstein, King, and West (2003) asserted that according to the social feedback theory, 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). This conclusion would support the social feedback theory. 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. Thus, additional empirical investigations are warranted.

Potentially Important Aspects of Vocalizations for Children with ASD

Various aspects of vocalizations can be used to describe vocal development and have potential validity for assessing vocal development in young children with ASD. We assert that volubility, communicative use, complexity, and reciprocity warrant further investigation based on several complementary theories as well as empirical evidence. After briefly defining each of these vocalization aspects here, in the following sections we describe theoretical and empirical evidence that provide the rationale for evaluating each aspect more thoroughly.

Volubility is defined as how frequently a child vocalizes (e.g., Patten et al., 2014). The specific types of vocalizations included in volubility measures have varied across researchers. For example, Shumway and Wetherby (2009) included only communicative vocalizations (i.e., excluded non-communicative vocalizations), and Oller et al. (2010) included only speech-related vocalizations (i.e., included prespeech vocalizations but excluded nonspeech-related sounds, such as crying and vegetative sounds). We define volubility as the number of vocalizations produced regardless of communicativeness or complexity (e.g., inclusion of a canonical syllable or consonant).

Communicative use of vocalizations is defined as how frequently or consistently a child produces vocalizations in an apparent attempt to transmit a message to another person (Wetherby, Yonclas & Bryan, 1989). Because children do not always direct their vocalizations to another person, one would expect the numerical value of a measure of communicative vocalizations to be less than the numerical value of a measure of total vocalizations.

Vocal complexity is defined as the frequency, consistency, or diversity with which a child produces vocalizations with certain features, such canonical syllables or consonants (e.g., Wetherby, Watt, Morgan, & Shumway, 2007; Woynaroski et al., 2017; Yoder, Watson, & Lambert, 2015). 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 the diversity of key consonants used in communication acts (DKCC; Wetherby et al., 2007; Woynaroski et al., 2017).

Vocal reciprocity is defined as the degree to which an adult vocal response to a child vocalization increases the likelihood of an immediately following child vocalization (Harbison et al., 2018). Vocal reciprocity measures capture the back and forth nature of vocal interactions.

Theoretical Support for Measuring Selected Aspects of Vocalizations

The theoretical bases for measuring various vocal variables overlap substantially. Although details vary across the social feedback theory, speech attunement framework, and transactional theory of spoken language development, these theories all center on the importance of interactions between children and adults for facilitating speech and language development. Thus, in addition to the older child-driven theoretical rationale for selecting the target vocal variables, one can make theoretical arguments involving bidirectional influence between child characteristics and adult input to justify the selection of volubility, communicative use of vocalizations, vocal complexity, and vocal reciprocity for facilitating language development in children with ASD. We describe below the key components of several potentially explanatory theories and hypotheses of language development that may be applied to children with ASD in general and to vocal development in this population more specifically.

Social feedback theory. As described above, Goldstein and colleagues (2003, 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. Similarly, Warlaumont, Richards, Gilkerson, and Oller (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 total and/or speech-like 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.

Speech attunement framework. "The speech attunement framework posits that the acquisition of articulate speech and appropriate prosody-voice requires a child to 'tune in' to the oral communications of the ambient community and to 'tune up' the phonological and phonetic behaviors subserving intelligible and socially appropriate speech, prosody, and voice production" (Shriberg, Paul, Black, & van Santen, 2011, p. 420). The speech attunement framework applies broadly to a number of speech production characteristics in children with ASD. In general, children with ASD may not "tune in" (i.e., attend to) and then be able to "tune up" to (i.e., broadly emulate) the general characteristics of adult speech due to deficits in social reciprocity. Deficits in self-monitoring speech, prosody, and voice also are posited to contribute to atypical vocal patterns in individuals with ASD (Shriberg et al., 2011). For young children with ASD in the early stages of language development, a reduced ability to "tune in" and "tune up" could explain their reduced use of typical, speech-like vocalizations.

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 & Yoder, 2002; Woynaroski, Yoder, Fey, & Warren, 2014). It posits that as a child's speech and language skills increase, parents provide more complex input that scaffolds continued child growth (Camarata & Yoder, 2002; Woynaroski et al., 2014). Despite some mixed findings, bidirectional influences have been documented for vocalization development in children with typical development (Fagan & 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 often and in what ways mothers respond to infant vocalizations (West & Rheingold, 1978; Yoder & Feagans, 1988). In initially preverbal children with ASD, Woynaroski et al. (2017) identified both parent (i.e., linguistic input) and child factors (i.e., intentional communication and receptive vocabulary) that predicted growth in DKCC. These findings provide support for the transactional theory of spoken language development in children with ASD. Additional support is provided by positive associations between parent verbal responsiveness (e.g., use of follow-in comments) and child spoken language skills in children with ASD (e.g., McDuffie & Yoder, 2010).

Application to vocal development in children with ASD. Theoretically, improving volubility, communicative use of vocalizations, vocal complexity, and/or vocal reciprocity could facilitate language development in children with ASD. Logically, increasing the frequency with which children with ASD vocalize (i.e., volubility) would increase the number of child vocalizations upon which the caregivers could respond. Additionally, children who vocalize frequently may provide themselves with opportunities to fine-tune their productions to their intended vocal targets through the auditory feedback loop mechanism (Koopmans van Benium, Clement, & Van Den Dikkenberg-Pot, 2001; Siegel, Pick, & Gerber, 1984). The social feedback theory, speech attunement framework, and transactional theory of spoken language development all emphasize the interactive roles that adults and children play in vocal development. Therefore, complexity, communicative use, and/or vocal reciprocity warrant attention. Increasing the complexity of children's vocalizations and/or the frequency with which children communicate with their vocalizations might elicit more frequent and/or more complex adult responses that then scaffold the child's ability to produce more adult-like productions including spoken words. Additionally, increases in communicative use or complexity of vocalizations could signal that children are attempting to say words that they understand but cannot yet produce accurately enough to be understood (Woynaroski et al., 2016). These theories of vocal development also provide theoretical support for the influence of vocal reciprocity. Increasing vocal reciprocity through intervention would increase the likelihood of adults responding to child vocalizations and children responding to adult vocal responses, which would increase the number of learning opportunities. Additionally, a relatively higher vocal reciprocity value may indicate that a child is attending to and affected by adult vocal responses. This relatively higher vocal reciprocity might increase the probability of the child up taking linguistic input from adults, which might facilitate the children's language development.

Empirical Support for Measuring Selected Aspects of Vocalizations

When establishing the construct validity of particular vocal variables for specific purposes, some of the most relevant pieces of evidence are the correlations between the vocal variables of interest with expressive language outcomes or measures of precursors to expressive language. These correlations would be most relevant to the current study if they came from studies of children with ASD in the early stages of language learning to provide an estimate of the association among vocal variables and expressive language in the ASD population. Broadly, a recent meta-analysis revealed that vocalizations correlate strongly with current or future expressive language skills for children with ASD (r = .50; 95% CI [.23, .76]; McDaniel, D'Ambrose Slaboch, & Yoder, 2018). Although the findings were stronger for concurrent associations (r = 0.77, 95% CI [0.45, 1.0]) than longitudinal associations (r = 0.33; 95% CI [0.05, 0.60]), the mean effect size for longitudinal associations was significant as well (McDaniel et al., 2018). Longitudinal associations provide stronger evidence of convergent validity than concurrent associations because in addition to providing evidence of an association between variables, they also provide evidence of a temporal precedence of the putative cause relative to the putative effect. Thus, longitudinal associations provide two of the three criteria for drawing a causal inference; whereas concurrent associations only provide one. The meta-analysis included a variety of vocal variables including those purported to measure volubility, communicative use, and vocal complexity. No studies reporting the correlation between vocal reciprocity at the prelinguistic level and expressive language for children with ASD met the inclusion criteria for the meta-analysis. In addition, the overall number of studies was too low to achieve sufficient power to test whether specific types of measures or variables yielded stronger correlations with expressive language than others. This correlational evidence from the meta-analysis provides support for continuing to evaluate potential relations between vocalizations and later expressive language in children with ASD.

Volubility. Correlations between volubility and expressive language for children with ASD have been reported. For example, the frequency of total vocalizations correlated at the same time with the Communication and Symbolic Behavior Scales (CSBS) Speech composite score, a measure of spoken expressive language and speech-like vocalizations, (r = .47; p < .01; Plumb & Wetherby, 2013) for 18- to 24-month-old children with ASD. The correlation between the frequency of total vocalizations during the second year of life and the verbal developmental quotient on the MSEL (Mullen, 1995), a measure of verbal impairment relative to chronological age, at age 3 was also significant for children with ASD (r = .39; p < .01; Plumb & Wetherby, 2013). In addition to being coded conventionally, volubility has been measured using the Language ENvironment Analysis (LENA) system (LENA Research Foundation, 2015). The number of child speech-related vocalizations correlated (r = .33) with age-equivalency scores on the Preschool Language Scale – Fourth Edition (Zimmerman et al., 2002) for 3- to 5-year-old children with ASD (Dykstra et al., 2013). In contrast, the number of child vocalizations per hour did not correlate significantly with expressive language measured concurrently by the Vineland Adaptive Behavior Scales (VABS; Sparrow, Cicchetti, & Balla, 2005) Expressive Language raw score (r = .10) or the MSEL Expressive Language raw score (r = -.24; Rankine, 2016) for a sample of children with ASD with a mean chronological age of 76.92 months (*SD* = 31.78

months). The current study evaluated the incremental validity of automatic versus conventional measures of volubility in predicting expressive language in children with ASD.

Communicative use. Correlations between communicative vocalizations specifically and expressive language in children with ASD have been reported as well. Plumb and Wetherby (2013) reported that communicative vocalizations in the second year of life predicted expressive language skills at age 3 above and beyond noncommunicative vocalizations. In contrast, Swineford (2011) did not report any significant correlations between the rate of communication acts with vocalizations within home observations and the CSBS Words subscale (r = .03) or the CSBS Speech composite (r = .13) concurrently for children suspected of having ASD (mean chronological age = 19.51 months; *SD* = 2.34 months).

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. The latter subordinate category of variables combines complexity concepts with communicative use concepts. Within this subordinate category of variables of vocal communication, the metric may be a count or a proportion in which the denominator refers to the communicative concept and the numerator refers to complexity concept. For example, several studies have examined consonant inventories within communicative vocalizations of children with ASD, rather than including all consonants regardless of communicative use. For the analyses and discussion purposes, we classify these variables within the complexity set of vocal variables because we judged the complexity component to be more prominent in the variable's interpretation. For example, DKCC is conceptually most related to consonant inventory in all vocalizations, which is clearly a complexity variable. Relatedly, proportion of communication acts with a canonical syllable is judged to be a complexity variable because it focuses on how consistently the child uses canonical syllables (a marker of complexity) as opposed to how consistently a child uses vocalizations for communicative purposes.

For use of consonants within vocalizations, Talbot (2014) reported that the rate of consonant-vowel vocalizations produced at 9 months of age correlated (r = .84) with expressive language on the MSEL at 12 months of age for children with ASD. The Systematic Observation of Red Flags for Autism Spectrum Disorders in Young Children (SORF) and the Systematic Observation of Red Flags for Autism Spectrum Disorders in Young Children at Home (SORF-Home) focus on the lack of communicative vocalizations with consonants as opposed to lack of vocalizations regardless of consonant use. MSEL verbal developmental quotient correlates negatively with concurrent lack of communicative vocalizations with consonants on the SORF (r = -.49; McCoy, 2013) and on the SORF-Home (r = -.57; Book, 2009). Relatedly, the rate of canonical babbling correlates with concurrent expressive language on the Reynell Developmental Learning Scales (r = .65, p < .05; Reynell & Gruber, 1990) in children with ASD (mean chronological age = 44.67 months; *SD* = 8.35 months; Sheinkopf, Mundy, Oller, & Steffens, 2000).

For consonant inventory measures, Yoder et al. (2015) found that the inventory of consonants used in communication acts predicted expressive language growth in initially preverbal children with ASD over and above ten other putative predictors. Similarly, Wetherby et al. (2007) identified that inventory of consonants used in communication acts at 18 to 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 (Woynaroski et al., 2017). In addition to the consonant inventory measures that consider communicative use of vocalizations, consonant inventory without regard to communicative use predicts expressive language skills. Consonant inventory size differed between children with ASD with functional communication versus children with ASD without functional communication (Paul, Chawarska, Cicchetti, & Volkmar, 2008).

Vocal reciprocity. Arguably, reasonable measures of vocal reciprocity are relatively new to the ASD literature. A new measure of vocal reciprocity, child reciprocal vocal contingency (RVC), warrants attention and continued investigation (Harbison et al., 2018). RVC quantifies vocal reciprocity from day-long audio samples collected in the child's natural environment. RVC uses the three-event sequence of a child vocalization followed immediately by an adult vocalization followed immediately by a child vocalization. Unlike one proposed measure of vocal turn-taking (i.e., child conversational turn count; Gilkerson & Richards, 2008) and another putative measure of vocal reciprocity using three events (Warlaumont et al., 2014), RVC is designed to control for the chance occurrence of child vocalizations and adult vocalizations mathematically. This mathematical control is an essential feature of a dyadic vocal reciprocity measure. For children with ASD who were preverbal or in the early stages of word learning, RVC correlated with consonant inventory used within communication acts (r = .60), but did not correlate with chronological age (r = -.001), intellectual quotient (r = -.23), or parents' formal education level (r = -.27; Harbison et al., 2018). Additionally, the association between RVC and consonant inventory in communication acts remained after statistically controlling for covariation with child volubility. These findings provide early evidence of construct validity for RVC and highlight the need for continued investigation of whether RVC measures what it purports to measure.

The Need for Additional Evidence of Validity for Vocal Variables

When comparing evidence of validity among multiple vocal variables, evidence is likely to show that some variables have multiple sources of evidence for validity and that other variables have less evidence of validity. The more sources of validity evidence, the more scientifically useful a vocal measure is likely to be. For example, a measure may exhibit strong evidence of convergent construct validity with expressive language, but weak evidence that it is sensitive to change. A single analysis or test is insufficient for reporting the degree to which a measure exhibits construct validity (Cronbach & Meehl, 1955). Instead, multiple sources of evidence must be integrated and evaluated for the specific purpose of the variable of interest.

Currently, nearly all evidence of validity for measuring vocal development in young children with ASD is convergent validity evidence (i.e., degree to which a variable correlates with other variables which it is predicted to correlate based on theory). Even though divergent validity is a key type of validity evidence (i.e., a variable does not correlate other variables with which it is not predicted to correlate based on theory; Campbell & Fiske, 1959), the literature base lacks such evidence for measuring vocal development in young children with ASD.

The field would benefit from comparisons of validity evidence across vocal variables purported to serve the same purpose. Meta-analytic approaches have been underpowered and difficult to interpret for comparing associations with expressive language due to a relatively low number of available primary studies (McDaniel et al., 2018). The current study presents an opportunity to compare directly convergent validity, divergent validity, sensitive to change validity, and incremental validity 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. Each vocal variable is described in detail in the Method section.

Comparing Validity Evidence Across Vocal Variables

Ultimately, this investigation sought to provide evidence of the comparative validity of different ways to quantify vocal development. However, there was no *a priori* consensus for how validity should be compared among competing variables. The research questions were organized around different ways to test the validity of variables. Within the discussion, two approaches organized the evidence to afford different methods of selecting vocal variables. The rationale for the approaches provided a rationale for the research questions.

The first approach presents the presence or absence of significance of association or change for evidence of convergent validity, divergent validity, and sensitivity to change. Some investigators criticize counting the number of significant findings as the basis for selecting among vocal variables because statistical significance relies on a method that is influenced by sample size and the concept of null hypothesis testing, which some have discredited (Krantz, 1999; Wilkinson, 1999). Therefore, we also used a threshold, above which we judged the effect size to be “large” (i.e., pseudo $R^2 \geq .25$ for convergent validity or Cohen’s $d \geq .8$ for sensitivity to change). Effect size criteria for selecting variables are not relevant for divergent validity. Additionally, the threshold values for “large” are only conventions and might not make sense in the context of evaluating relative validity of vocal variables. Comparing significance and effect size across all vocal variables for all purposes simplifies the task of selecting vocal variables. However, selecting the variables with the greatest number of significant or large effect sizes across purposes ignores the fact that the most useful variable for a particular purpose may fail to provide value for other purposes. Arguably, a more useful approach is to select the vocal variables with significant and large effect sizes within a purpose (e.g., to predict expressive language or to show evidence of change). However, this comparison method ignores the fact that a slightly less valid measure may cost much less than a slightly more valid measure.

Thus, incremental validity of predicting expressive language for more elaborate or more costly variables was compared with that of simpler or more cost-effective variables as the second approach. Using variables that are more elaborate and/or costly 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. Three ways to conceptualize elaborateness or cost are used in the current research. First, we compared simple volubility variables with more elaborate (and if measured from conventional communication samples, more costly) communicative use and complexity variables. Second, we compared RVC, which is a dyadic three-event variable, with communicative use and complexity variables, which are single actor, single event variables. Third, we compared the less costly data collection and variable derivation approach of automatic measures with the more costly approach of conventionally-coded variables, which requires human segmenting and classification of child vocalizations. To make elaborate (more costly) versus simple (less costly) variable comparisons, we used the concept and methods of testing incremental validity (i.e., significant association with expressive language after statistically controlling for another variable) to identify whether the more elaborate or costly measure added information above and beyond the information provided by the less elaborate or less costly measure.

Purpose and Research Questions

This study was designed to evaluate and compare quantitative evidence of validity for vocal variables purported to assess vocal development in young children with ASD. We addressed each of the following research questions for vocal variables purported to assess the following aspects of vocalizations in this population: (a) volubility, (b) communicative use of vocalizations, (c) vocal complexity, and (d) vocal reciprocity. Within volubility and vocal complexity, there are variables derived by automated (i.e., collected by LENA recording devices and computer analyzed without human transcription) and conventional coding (i.e., communication sampling methods and human coding of vocalizations) methods. Research questions 4, 5, and 6 use composite variables when theoretically and empirically justified to provide the best estimate possible for volubility, communicative use of vocalizations, and vocal complexity. One vocalization aspect, vocal reciprocity, cannot be measured using a composite because only one known measure meets content validity for the construct.

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?
4. Compared with the relatively simple volubility vocal variables, does communicative use, complexity, or reciprocity account for additional unique variance in expressive language skills?
5. Compared with the single actor, single event (less elaborate) communicative use and complexity variables, does the dyadic three-event (more elaborate) measure of vocal reciprocity account for unique variance in expressive language skills?
6. Compared with less costly automated measures of the same vocalization aspect, do more costly conventionally-coded variables account for unique variance in expressive language skills?

CHAPTER 2

METHOD

Institutional Review Boards at the Vanderbilt University, the University of Washington, and the University of Southern California at Davis approved all study procedures. Caregivers provided written informed consent prior to participants beginning the study.

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, Estes, & Yoder, 2013). The TADPOLE study compared language and developmental outcomes of a sample of young children with ASD who were randomly assigned to a treatment style (i.e., discrete trial training or play-based using the Early Start Denver Model [Rogers & Dawson, 2009]) and intensity (i.e., 15 or 25 hours per week). More details on the intensity and style manipulations can be found in the grant application (Rogers et al., 2013). The details are not given here because there were few interactions with style or intensity. When such interactions occur, we simply examine associations within the relevant treatment groups due to an agreement with the principal investigator of the parent grant. The focus of this dissertation study is relative validity of vocal variables, not treatment efficacy. There is no emphasis on interpreting treatment effects in the current study. 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 diagnostic criteria for ASD, (d) overall developmental quotient of at least 35 on the MSEL (mental age / chronological age x 100), (e) English as a primary language (i.e., English spoken at least 60% of the time at home per caregiver report), and (f) hearing and visual acuity within normal limits per screening results.

ASD diagnosis was based on all of the following: (a) Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition diagnostic criteria for ASD (American Psychiatric Association, 2013), (b) clinical consensus of diagnosis by two independent staff, one of whom is a licensed psychologist, based on observation and record review, (c) meeting full autism criteria on the Autism Diagnostic Interview-Revised (Lord, Rutter, & Le Couteur, 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 assessor who evaluated the child.

For study eligibility, the participants' caregivers also agreed to complete the project, including parent coaching at home, in-home therapy 15 or 25 hours per week, and monthly clinic visits. Participants were excluded if they had been receiving 10 or more hours per week of intensive, curriculum-based therapy for at least one month. Participants were not excluded based on the presence of genetic disorders or other health conditions (e.g., Fragile X syndrome, seizures, and prematurity) in addition to ASD.

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.

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

Note. 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).

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 / initiation of intervention; Time 2 = 6 months post study initiation / intervention midpoint; Time 3 = 12 months post study initiation / end of intervention). As described below some procedures were not administered at Time 2 due to resource constraints in the parent grant's budget.

Table 2

Study Constructs, Procedures, and Variables

Construct	Procedure(s)	Variable
Volubility	CSP & ECI	Number of total vocalizations
	Day-long audio samples	Number of child speech-related vocalizations
Communicative use of vocalizations	CSP & ECI	Number of CAs that include a vocalization
		Number of CAs that include a canonical syllable
		Proportion of vocalizations that are communicative
Vocal complexity	CSP & ECI	Consonant inventory (regardless of communicative use)
		DKCC
		Proportion of CAs with a canonical syllable
		Proportion of vocalizations with a canonical syllable
		Number of vocalizations with a canonical syllable
	Day-long audio samples	ACPU-Consonants
		ACPU-Vowels
		IVD score
Vocal reciprocity	Day-long audio samples	RVC
Expressive Language	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

Note. ACPU = Average Count Per Utterance (Xu, Richards, & Gilkerson, 2014); CAs = communication acts; CSP = Communication Sample Procedure; DKCC = diversity of key consonants used in communication acts (Wetherby, Watt, Morgan, & Shumway, 2007; Woynaroski et al., 2017); ECI = Early Communication Index (Greenwood, Carta, Walker, Hughes, & Weathers, 2006; Luze, Linebarger, Greenwoord, & Carta, 2001); IVD = infraphonological vocal development (Oller et al., 2010); MB-CDI = MacArthur-Bates Communicative Development Inventory: Expressive vocabulary compilation form (Fenson et al., 2007); MSEL = Mullen Scales of Early Learning (Mullen, 1995); RVC = reciprocal vocal contingency (Harbison et al., 2018); VABS = Vineland Adaptive Behavior Scales, Second Edition (Sparrow, Cicchetti, & Balla, 2005).

Day-long naturalistic audio samples. Participants' families collected one day-long audio recording at Time 1 and one at Time 3 with the LENA system (LENA Research Foundation, 2015). The research team provided families with all necessary equipment and instructions. The LENA digital recording device (i.e., digital language processor [DLP]) was placed in the pocket of a specialized vest for the participant to wear for a full day (at least 12 hours) with a goal of 16 hours of recording. No specifications were given regarding the day of the week or setting for the recording, except that the participant should not be ill or going swimming or to the pool with the LENA recording device. Caregivers were instructed to remove the participant's vest, with the recorder still on, and place it near the participant when he or she was sleeping or in the car. Families returned the DLP and other equipment to the research team. Trained research assistants downloaded the digital audio files from the DLPs to a designated computer for processing and analysis.

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. The CSP was administered at Times 1 and 3.

Early Communication Index (Greenwood, Carta, Walker, Hughes, & Weathers, 2006; Luze, Linebarger, Greenwood, & Carta, 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 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 intervention period (i.e., Time 1 to Time 3) for a total of 13 administrations. To align with the available CSP data, we averaged the ECI sessions from the first 3 months for Time 1 and the sessions from the last 3 months for Time 3.

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 two 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, Cicchetti, & Balla, 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.

Observational Coding

Trained research assistants and the first author completed observational coding for the CSP and ECI using ProCoder DV (Tapp, 2003) and Systematic Analysis of Language Transcripts (SALT) software (Miller & Chapman, 2016). Volubility, production of vocalizations in communication acts, vocal complexity, and number of different words were coded from the CSP and ECI. CSP and ECI session variables from the same time period were averaged after checking for a sufficiently high correlation between them (i.e., $r \geq .40$).

Coders completed four passes using timed event behavior sampling to code behaviors in the CSP and ECI necessary for deriving the vocal development variables. Trained research assistants on our team had already completed the first two passes. On the first pass, the coder identified codable and uncodable portions of each video file. "Uncodable" is defined as a period of at least 10 seconds during which the child's face is not visible (e.g., turned away from camera or blocked from view by another person or an object). On the second pass, the coder identified all communication acts within the codable time. The coder also classified each communication act as symbolic (i.e., single non-imitated word or multiple non-imitated words) or non-symbolic (i.e., imitated words or phrases, non-word vocalizations, and gestures) and orthographically transcribed words that the child said. The coding manual includes detailed communication act coding rules. See Table 3 for operational definitions of key concepts for coding.

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 seconds 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 sounds (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: • 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

The third and fourth passes were completed for the current study. Within the third pass, the coder identified and classified vocalizations that occurred within a communication act to indicate whether they contain 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. On the final pass, the coder listened to the entire recording stopping it each time he or she heard a vocalization. For any vocalizations not already coded as part of a communication act, the coder marks the vocalization as a non-communicative vocalization and indicates whether it includes one or more codable consonants and/or a canonical syllable.

Variables Derived from Observational Coding

SALT software was used to calculate variables for each vocalization aspect when the variable was derived from the ECI or CSP. See Table 2 for the study variables.

Volubility variables. Volubility is defined as the number of vocalizations produced. The vocalizations must be nonvegetative voiced sounds produced during exhalation. This variable includes communicative and noncommunicative vocalizations. One of the two volubility variables was derived from observational coding: number of total vocalizations.

Communicative use of vocalizations variables. A communication act is defined as an intentional behavior or set of intentional behaviors that meets criteria for one of the following: (a) spoken or signed word(s), (b) nonword vocalization(s) with evidence of coordinated attention, or (c) one of 15 specific gestures with evidence of coordinated attention to message/referent and communication partner. Three variables were calculated for vocalizations in communication acts: (a) the number of communication acts that include a vocalization, (b) the number of communication acts that include a canonical syllable, and (c) the proportion of vocalizations that are communicative (i.e., within a communication act). Redundant variables were eliminated during the preliminary analyses based on an *a priori* intercorrelation level (i.e., $r = .80$).

Vocal complexity variables. A relatively large number of vocalization complexity variables were derived due to gaps in the literature regarding how to measure individual differences and changes in vocal complexity. Consonant inventory (regardless of communicative use), DKCC (Wetherby et al., 2007; Woynaroski et al., 2017), the proportion of communication acts with a canonical syllable, the proportion of vocalizations with a canonical syllable (regardless of communicative use), and the number of vocalizations with a canonical syllable were coded from the CSP and ECI. The proportion as well as the number of vocalizations with canonical syllables was coded because, theoretically, the consistency of canonical syllable use and number of times a child uses canonical syllables could be related to expressive language development, but for different reasons. As with the communicative use variables, redundant vocal complexity variables were eliminated empirically during the preliminary analyses.

Automated Vocal Analysis

To derive the five variables from the day-long audio recordings (see Table 2), we used several methods. The number of speech-related child vocalizations is available through the standard LENA Pro software package. However, Average Count Per Utterance – Consonants (ACPU-C; Xu, Richards, & Gilkerson, 2014), Average Count Per Utterance – Vowels (ACPU-V; Xu et al., 2014), infraphonological vocal development (IVD) score (Oller, et al., 2010), and RVC (Harbison, et al., 2018) scores are not yet available in the standard or research LENA software packages. The ACPU-C, ACPU-V, and IVD scores require raw scores derived the audio recordings via computer programs housed at the LENA Research Foundation. These variables were available as a result of a contract between Dr. Paul Yoder and the LENA Research Foundation. RVC was calculated through a freely available and publicly posted software program that uses the LENA Integrated Time Segments (ITS) files (Xu, Yapanel, Gray, & Baer, 2008) as input (Harbison et al., 2018).

Regardless of the process through which each variable was obtained, all of the study's automated vocal analyses rely on how the LENA system segments acoustic events and determines the sound source for each sound segment in the recordings. In the first step, acoustic events from the audio recording are divided into short segments. In the second step, the short sound segments are classified into one of eight preliminary categories: key child (i.e., child wearing the audio recorder), other child, adult male, adult female, overlapping sound, television and other electronic sound, noise, and silence. In the third step, the fit of each segment to the segment's preliminary classification is compared with the silence model. If the segment is different from the silence model, the preliminary classification is maintained (e.g., "key child" remains "key child"). If not, the segment is reclassified as a secondary classification with a "faint" notation (e.g., "key child" classification becomes "key child – faint" classification). At this point, each sound segment has been identified and classified as one of 15 sound source categories.

Sound segments that retained their classification as key child are analyzed further using a six-step process. First, the program divides key child events into one or more "child utterance clusters" (CUCs). The CUCs have key child-near as their source, are at least 600 ms, and are not interrupted by another speaker or by more than 800 ms of silence or noise. Second, "child utterances" (CUs) within the CUCs are identified. CUs are related to breath-groups of child vocalizations and interrupted for no more than 300 ms. Third, "child vocal islands" (CVIs) are identified in the CUs by patterns of high energy relative to the baseline energy level. Fourth, the CVIs are classified as (a) cries, (b) vegetative sounds (e.g., laughs, sneezes, or coughs), or (c) "speech-related vocal islands" (SVIs). The concept of SVIs is related to but not synonymous with syllables. Fifth, the SVIs are grouped or lumped into speech-related child utterances (SCUs), which are interrupted by sounds from other sources that are no more than 300 ms in duration. SCUs include words, babbling, and protophones (e.g., squeals, growls, and raspberries; Xu et al., 2008). Each SCU is considered analogous to a single child vocalization

(Oller et al., 2010). Adult vocalizations are identified through an analogous process. See Xu, Yapanel, and Gray (2009), VanDam and Silbert (2016), and Rankine (2016) for information on the reliability of LENA system's classification of the audio recordings relative to human coding. The ITS file for each recording includes the stream of events (e.g., child vocalizations, adult vocalizations, and silence) and is used for various analyses.

Variables Derived from Automated Vocal Analysis

Volubility variable. For the automated vocal analysis measure, volubility is operationalized as the number of near child speech-related vocalizations (i.e., SCUs). These vocalizations are identified as SCUs within the LENA system (Oller et al., 2010). We generated this count variable of the number of child speech-related vocalizations when calculating RVC as described below. To be as similar as possible to the conventional measure of volubility, we used the total number of near child speech-related vocalization across the entire day-long recording. We did not use a rate metric that is available through the standard LENA software (i.e., number of near child vocalizations per hour, day, or month; Gilkerson & Richards, 2008).

Vocal complexity variables. IVD score is conceptualized as a measure of vocal complexity (Oller et al., 2010; Yoder, Oller, Richards, Gray, & Gilkerson, 2013). To calculate IVD score, the SVIs within each SCU were classified as either possessing or not possessing each of 12 characteristics of speech-likeness/syllabicity. IVD score is calculated the weighted sum of the raw scores. The raw scores are proportions of SVIs with a particular speech-like characteristic. The weights in the sum are the unstandardized regression coefficients from a multiple regression equation predicting age in the Oller et al. (2010) normative sample. The LENA Research Foundation provided the raw scores and the weights, and we computed IVD scores.

The ACPU-C and ACPU-V scores are based on Sphinx speech recognition software. Sphinx is used to derive the ACPU scores by estimating the number of times 39 phones occur within utterances produced in SCUs (Xu et al., 2014). The LENA Research Foundation uses noncommercially-available software that implements an algorithm that considers child age and gender information. It should be noted that Sphinx software was modeled only with adult data. For information about the validation process see Xu et al. (2008). ACPU-C and ACPU-V are related conceptually. Average Count Per Utterance – Consonants and Vowels (ACPU-C+V) is created by aggregating *z*-scores from ACPU-C and ACPU-V scores (Woynaroski et al., 2017) when the ACPU component variables correlate at or above $r = .40$.

Vocal reciprocity variable. RVC is the operant contingency value for a three-event sequence: child vocalization (CV) → adult vocalization (AV) → CV (Harbison et al., 2018). From a content validity perspective, the back and forth nature of the vocal reciprocity construct arguably necessitates use of a three-event sequence. The RVC program uses the event lag with contiguous pauses sequential analysis method to prepare the vocal samples for analysis. This method retains the events of interest (i.e., CVs and AVs) and removes other events while maintaining the event sequence and temporal proximity of events by inserting fixed-duration pause events when neither key event occurs (Lloyd, Yoder, Tapp, & Staubitz, 2016). A simulation study found this method to be more accurate and less correlated with chance occurrence of the events of interest than other sequential analysis methods (Lloyd et al., 2016). We inserted 2-s pauses based on the duration of pauses in conversations between adults and infants with typical development and high-risk infant siblings of children with ASD (Gros-Louis et al., 2006; Northrup & Iverson, 2015). The three-event sequences composed of various combinations of the CVs, AVs, and 2-s pauses were tallied into one of the four cells in a 2x2 contingency table, depending on the sequence and occurrence of the event types. Using the cell labels in Figure 1, the RVC value was computed using the formula [a/(a+b)] - [c/(c+d)].

Positive RVC scores provide correlational evidence that immediately preceding adult vocal responses to preceding child vocalizations influence the child's following vocal response. A positive RVC indicates that child vocalizations are more likely to follow adult vocal responses to child vocalizations than other events. Importantly, compared to other proposed measures of sequential association, the operant contingency value mathematically controls for the chance sequencing of the events as well or better than others (Hammond, 1980; Lloyd, Kennedy, & Yoder, 2013; Martens, Gertz, Werder, Rymanowski, & Shankar, 2014). RVC has been shown to be more related to precursors of expressive language in initially preverbal or early verbal children with ASD than competing measures of vocal reciprocity, particularly when controlling for chance sequencing of events (Harbison et al., 2018).

Figure 1

Contingency Table for Calculating Reciprocal Vocal Contingency

		Event 3	
		CV	(not CV)
Events 1 and 2	[CV → AV]	a [CV → AV] → CV	*b* [CV → AV] → (not CV)
	(not [CV → AV])	*c* (not [CV → AV]) → CV	*d* (not [CV → AV]) → (not CV)

Note. 2x2 contingency table; cell labels (i.e., a, b, c, and d) are centered at the top of each of the four cells; AV = adult vocalization; CV = child vocalization; CV → AV = child vocalization followed by an adult vocalization within 2 s without any intervening events.

Interobserver Reliability

A trained secondary coder independently coded a random sample of $\geq$ 20% of coded sessions for each time point for variables derived from the CSP and ECI. 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 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 three consecutive videos (Yoder, Lloyd, & Symons, 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.

CHAPTER 3

RESULTS

Preliminary Analyses

Before addressing the primary research questions, we assessed each variable's reliability, eliminated redundant vocal variables, and created several types of composite variables for specific purposes. Composite variables created by aggregating multiple variables increase the short-term stability of constructs of interest (Sandbank & Yoder, 2014; Yoder et al., 2018). Theoretically-related variables were only aggregated if they exhibited a correlation coefficient of $r \geq .40$ (Cohen & Cohen, 1984).

Reliability. For all conventionally-coded variables combined, the mean ICC was .93 (*SD* = .11). ICCs are reported for the conventionally-coded variables by time period and procedure in Table 4. 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 3 for Time 3. We used a benchmark of .70 when interpreting the ICCs, which Mitchell (1979) interpreted as "very good". Only two variables fell below .70 for the ICC: ECI Month 3 proportion of vocalizations that are communicative (ICC = .64) and CSP Time 3 proportion of communication acts with a canonical syllable (ICC = .24). ECI Month 3 proportion of vocalizations that are communicative was included in the analyses because although .64 is below .70, ECI Time 3 proportion of vocalizations that are communicative correlated strongly with the same variable measured within the CSP at Time 3 (r = .75). This high correlation suggests that the ECI Time 3 proportion of vocalizations that are communicative is likely to be useful scientifically. In contrast, CSP Time 3 proportion of communication acts with a canonical syllable exhibited a low correlation with the same variable within the ECI at Time 3 (r = .31). This low correlation provides additional evidence that CSP Time 3 proportion of communication acts with a canonical syllable is not sufficiently reliable to be useful scientifically. Thus, we excluded this variable at both time points from all analyses.

Table 4

Intraclass Correlation Coefficients for Conventionally-Coded Vocal Variables by Time and Procedure

	Time 1		Time 3	
Variable	CSP	ECI Mean (*SD*)	CSP	ECI Mean (*SD*)
Number of total vocalizations	.99	.99 (.01)	.99	.99 (.01)
Number of CAs that include a vocalization	.97	.97 (.01)	.99	.98 (.02)
Number of CAs that include a canonical syllable	.97	.98 (.02)	.99	.97 (.03)
Proportion of vocalizations that are communicative	.84	.84 (.17)	.98	.96 (.03)
Consonant inventory (regardless of communicative use)	.94	.89 (.04)	.93	.96 (.01)
DKCC	.79	.90 (.02)	.98	.96 (.02)
Proportion of CAs with a canonical syllable	.89	.81 (.07)	.24	.87 (.07)
Proportion of vocalizations with a canonical syllable	.96	.92 (.06)	.97	.92 (.04)
Number of vocalizations with a canonical syllable	.97	.98 (.01)	.99	.98 (.02)

Note. CSP = Communication Sample Procedure; DKCC = diversity of key consonants used in communication acts (Wetherby, Watt, Morgan, & Shumway, 2007; Woynaroski et al., 2017); ECI = Early Communication Index (Greenwood, Carta, Walker, Hughes, & Weathers, 2006; Luze, Linebarger, Greenwoord, & Carta, 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.

Eliminating redundant vocal variables. We identified and removed variables highly correlated with other variables that purport to measure the same aspect of vocalizations at Time 1 because such high correlations indicate redundancy. The redundant variables were excluded from all analyses (i.e., Times 1 and 3). For communicative use, the communication samples were coded for (a) the number of communication acts that include a vocalization, (b) the number of communication acts that include a canonical syllable, and (c) the proportion of vocalizations that are communicative. The number of communication acts with a vocalization correlated almost perfectly with the number of communication acts with a canonical syllable at Time 1

(r = .99). See Appendix A for intercorrelations between vocal variables of the same vocalization aspect. We retained communication acts that include a vocalization to provide a variable varying from the total vocalizations volubility variable in only one aspect (i.e., communicativeness). Thus, we eliminated the number of communication acts that include a canonical syllable.

For complexity, the communication samples were coded for five variables (see Table 2). Recall that the proportion of communication acts with a canonical syllable was eliminated due to low reliability. Consonant inventory (regardless of communicative use) and DKCC correlated very highly at Time 1 (r = .85). Number of vocalizations with a canonical syllable also correlated highly with DKCC (r = .86) at Time 1. In children with ASD, there is relatively greater empirical evidence for the association between expressive language and DKCC compared with the other vocal variables (e.g., Wetherby et al., 2007; Yoder et al., 2015). Thus, we eliminated consonant inventory and the number of vocalizations with a canonical syllable. DKCC and the proportion of vocalizations with a canonical syllable remained.

Creating composite variables. If the intercorrelation among component variables posited to measure the same construct warranted it, we created composite variables for the expressive language and the nonverbal cognitive skills constructs to be used in the multilevel models, for variables derived from conventional coding of the CSP and ECI, and for the complexity variables derived from automated analyses. We calculated and averaged the *z*-scores for each component variable using the sample's mean and standard deviation at Time 3 to create the composite variable of interest.

The component variables of the expressive language composite (see Table 2) all correlated with each other at $r \geq .40$ at each time point. The component variables of the nonverbal cognitive skills composite (see Table 2) all correlated 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 each time point, both were retained for the composite at Time 2.

We created composite variables for variables conventionally-coded from the CSP and ECI. ECI values were calculated by averaging values across months 1 through 3 for Time 1 and months 11 through 13 for Time 3. Thus, values for each conventional coding variable include data from one CSP sample and three ECI samples. Correlations between Time 1 CSP and Time 1 ECI correlated at $r \geq .60$ for all of the retained vocal variables. The *z*-scores for the CSP and ECI values were averaged to create variables for the conventionally-coded measures.

To reduce the number of analyses used to address the incremental validity research questions, we computed composites for retained component variables at Time 1 if they correlated above $r = .40$. The number of communication acts with a vocalization and the proportion of vocalizations that are communicative correlated sufficiently ($r = .70$). Similarly, DKCC and proportion of vocalizations with a canonical syllable also correlated sufficiently ($r = .79$). The average *z*-score transformations of the component variables were the composite variables.

We planned to create a composite variable from the ACPU-C, ACPU-V, and IVD score automated variables for increased stability and construct validity (Woynaroski et al., 2017). Because ACPU-C and ACPU-V correlated strongly ($r = .82$), we aggregated them to form the ACPU-C+V variable by calculating and averaging the *z*-scores for each component variable. However, ACPU-C+V was not sufficiently correlated with IVD score ($r = .26$). Thus, ACPU-C+V and IVD score were analyzed separately for all analyses.

For volubility, total vocalizations measured via conventional coding and the number of child speech-related vocalizations from automated analyses were not sufficiently correlated to create a composite variable ($r = .37$ at Time 1; $r = .27$ at Time 3). Thus, total vocalizations and number of child speech-related vocalizations were analyzed separately.

Evaluating Evidence of Convergent Validity

For evidence of convergent validity we tested whether each vocal variable predicted later expressive language 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, which is the end of the intervention 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 mixed level modeling is to identify the unconditional growth model. We used a build-up approach for model selection. The random intercept, fixed slope model provided evidence of a better fit than a fixed intercept, fixed slope model (i.e., empty model). The -2 log likelihood value decreased from 633 for the fixed intercept, fixed slope model to 421 for the random intercept, fixed slope model. Although the -2 log likelihood value decreased further for the random intercept, random slope model relative to the random intercept, fixed slope model, 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 high covariance between the intercept and slope means that there is limited variance remaining to be explained by predictor variables in the model. The growth parameter of interest was intercept, which was interpreted as the best estimate of end point (Time 3) expressive language.

The style and/or intensity of treatment that the participants received could potentially influence the correlations among key variables. That is, it is possible that the strength of the association with later expressive language and/or patterns of change differ for vocalization measures for children in different treatment groups. Because the sample size prevented reasonable analysis of intensity x style x predictor three-way interactions, we examined whether the association between vocal variables and the intercept of growth of expressive language varied by intensity or style. To check for interactions, we added a group by predictor product term to the main effects of group and predictor in the model. None of the product terms between group and predictor were significant for the models predicting expressive language. Consequently, the models for this research question include participants pooled across groups.

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 the Time 3 expressive language) one at a time. As shown in Table 5, all variables were significant predictors except for IVD score. All significant associations were positive. See Appendix B for complete results for all models. No evidence of heteroscedasticity was observed. All residuals fell within the acceptable parameters for skewness and kurtosis of $< |.8|$ and $< |3.0|$, respectively (Tabachnick & Fidell, 2001). The pseudo R^2 value provides an effect size that represents the amount of explainable variance explained by the predictor variable (e.g., number of total vocalizations). It is the proportional reduction in residual variance that occurs when a predictor is added (Singer & Willett, 2003). Mathematically pseudo R^2 is the difference between the residual variance of the intercept between the full model (i.e., includes the vocal predictor variable of interest) and the reduced model (i.e., excludes the vocal predictor variable of interest) divided by the residual variance of the intercept for the reduced model. Conceptually, pseudo R^2 means the proportion of the growth model or growth parameter that is 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, number of communication acts that include a vocalization, proportion of vocalizations that are communicative, DKCC, and proportion of vocalizations with a canonical syllable at Time 1 have a large association with the best estimate of later expressive language.

Table 5

Fixed Effects Estimates for Vocal Variables Predicting End Point Expressive Language

Vocal Variable	Coeff.	*SE*	*t*	*df*	*p*	Pseudo R^2
Number of total vocalizations	0.26	0.07	3.74	88.57	<.001	.16
Number of child speech-related vocalizations	$2.9\text{x}10^{4}$	$6.8\text{x}10^{5}$	4.35	83.89	<.001	.19
Number of CAs 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
DKCC	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
ACPU-C+V	0.27	0.10	2.71	86.50	.01	.08
IVD score	0.00	0.01	-0.32	84.66	.75	.00
RVC	2.45	0.92	2.67	85.67	.01	.08

Note. ACPU-C+V = Average count per utterance – consonants + vowels (Woynaroski et al., 2017; Xu, Richards, & Gilkerson, 2014); Coeff. = coefficient value; DKCC = diversity of key consonants used in communication acts (Wetherby, Watt, Morgan, & Shumway, 2007; Woynaroski et al., 2017); IVD = Infraphonological vocal development (Oller et al., 2010); Pseudo R^2 = 1 – (covariance parameter for intercept of model with the vocal variable of interest / covariance parameter for intercept of model without the vocal variable of interest); RVC = reciprocal vocal contingency (Harbison et al., 2018).

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, which can be interpreted as predicting the best estimate of end point (i.e., Time 3) nonverbal cognitive skills. Theoretically, vocalization measures should *not* be significant predictors of nonverbal cognitive skills. 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. We rely on significance to define what is considered low. Given the large sample size, this approach is quite reasonable.

As with the model predicting expressive language, we used a build-up approach for model selection for the model predicting nonverbal cognitive skills. The random intercept, random slope model provided evidence of a better fit than the fixed intercept, fixed slope model and the fixed intercept, random slope model. The -2 log likelihood value decreased from 633 for the fixed intercept, fixed slope model to 421 for the random intercept, fixed slope model to 314 for the random intercept, random slope model. The intercorrelation between the intercept and slope for the random intercept, random slope model was acceptable (r = .79). As shown in Table 6, only the main effect for total vocalizations was a significant predictor of end point nonverbal cognitive skills. See Appendix C for complete results for all models. No evidence of heteroscedasticity was observed. All residuals fell within the acceptable parameters for skewness and kurtosis (Tabachnick & Fidell, 2001).

Table 6

Fixed Effects Estimates for Main Effects of Vocal Variables Predicting End Point Nonverbal Cognitive Skills

Vocal Variable	Coeff.	*SE*	*t*	*df*	***p***
Number of total vocalizations	0.13	0.06	2.16	86.42	.03
Number of child speech-related vocalizations	6.7×10^{5}	5.8×10^{5}	1.16	84.49	.25
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
DKCC	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
ACPU-C+V	-0.10	0.08	-1.24	84.85	.22
IVD score	0.01	0.01	1.04	84.53	.30
RVC	0.13	0.73	0.17	84.55	.86

Note. ACPU-C+V = Average count per utterance – consonants + vowels (Woynaroski et al., 2017; Xu, Richards, & Gilkerson, 2014); Coeff. = coefficient value; DKCC = diversity of key consonants used in communication acts (Wetherby, Watt, Morgan, & Shumway, 2007; Woynaroski et al., 2017); IVD = infraphonological vocal development (Oller et al., 2010); RVC = reciprocal vocal contingency (Harbison et al., 2018).

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 except IVD score exhibited evidence of sensitivity to change. The only significant group by predictor interaction was for number of child speech-related vocalizations. The high intensity group demonstrated evidence of sensitivity to change ($t(36) = 4.25$, $p < .001$). The low intensity group did not ($t(36) = -0.77$, $p = .45$). There was no significant difference between the high and low intensity groups at Time 1 for number of child speech-related vocalizations ($t(82) = 0.93$, $p = .35$).

Table 7

Results of Paired t-Tests from Time 1 to Time 3

Vocal Variable	Mean	*SD*	95% CI	*t*	*d*
Number of total vocalizations	0.76	0.76	[0.59, 0.93]	8.95	0.88***
Number of child speech-related vocalizations	410.87	1074.74	[161.87, 659.86]	3.29	0.40**
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***
DKCC	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***
ACPU-C+V	0.71	0.96	[0.49, 0.93]	6.30	0.85***
IVD score	2.52	11.59	[-0.19, 5.22]	1.86	0.27
RVC	0.02	0.09	[0.00, 0.04]	1.96	0.27*

Note. * = $p < .05$, ** = $p < .01$; *** = $p < .001$; *p* values are for two-tailed significance tests; ACPU-C+V = Average count per utterance – consonants + vowels score (Woynaroski et al., 2017; Xu, Richards, & Gilkerson, 2014); *d* = within subjects effect size accounting for correlation between Time 1 and Time 3; DKCC = diversity of key consonants used in communication acts (Wetherby, Watt, Morgan, & Shumway, 2007; Woynaroski et al., 2017); IVD = infraphonological vocal development (Oller et al., 2010); RVC = reciprocal vocal contingency (Harbison et al., 2018).

Evaluating Incremental Validity Relative to Volubility

To evaluate the incremental validity for aspects of vocalizations, we assessed the incremental validity of composite measures of these vocalization aspects when possible. Recall that the composite for communicative use includes (a) the number of communication acts that include a vocalization and (b) the proportion of vocalizations that are communicative. Similarly, the composite for complexity includes (a) DKCC and (b) the proportion of vocalizations with a canonical syllable. Based on the evidence for convergent validity, divergent validity, and sensitivity to change, we used ACPU-C+V instead of IVD score to evaluate the incremental validity of an automated variable for vocal complexity.

To test for incremental validity relative to volubility, we added communicative use, complexity, or reciprocity vocal variables to a model with volubility. Because total vocalizations and the number of child speech-related vocalizations were not sufficiently correlated, separate models were used for each. No predictor by group interaction effects were observed. The unstandardized coefficients, standard errors, and significance for the predictor variables are displayed in Tables 8 and 9. See Appendix D for complete results for all models.

Table 8

Unstandardized Coefficients (Standard Errors) and Significance for More Elaborate Vocal Variables Predicting End Point Expressive Language After Controlling for Conventionally-Coded Volubility

	TV	Comm. composite	Complexity composite	ACPU-C+V	RVC	Pseudo R^2 change
Comm.	-0.01 (0.06)	0.91 (0.10)***				.61
Complexity	-0.11 (0.07)		0.60 (0.08)***			.50
	0.27 (0.07)***			0.24 (0.09)**		.10
RVC	0.25 (0.07)**				1.79 (0.88)*	.07

Note. * = $p < .05$; ** = $p < .01$; *** = $p < .001$; ACPU-C+V = Average count per utterance – consonants + vowels score (Woynaroski et al., 2017; Xu, Richards, & Gilkerson, 2014); Comm. = communicative use; Pseudo R^2 change = 1 – (covariance parameter for intercept of model with more elaborate vocal variable added / covariance parameter for intercept of model only including volubility); RVC = reciprocal vocal contingency (Harbison et al., 2018); TV = total vocalizations.

Table 9

Unstandardized Coefficients (Standard Errors) and Significance for More Elaborate Vocal Variables Predicting End Point Expressive Language After Controlling for an Automated Volubility Variable

	CHNSP	Comm. composite	Complexity composite	ACPU-C+V	RVC	Pseudo R^2 change
Comm.	$9.7\text{x}10^5$ ($5.2\text{x}10^5$)	0.83(0.09)***				.62
Complexity	$9.5\text{x}10^5$($6.9\text{x}10^5$)		0.47(0.06)***			.49
	$2.6\text{x}10^4$($6.9\text{x}10^5$)***			0.17(0.09)		.05
RVC	$2.6\text{x}10^4$($7.9\text{x}10^5$)**				0.79(1.00)	.01

Note. ** = p < .01; *** = p < .001; ACPU-C+V = Average count per utterance – consonants + vowels score (Woynaroski et al., 2017; Xu, Richards, & Gilkerson, 2014); CHNSP = number of child speech-related vocalizations; Comm. = communicative use; Pseudo R^2 change = 1 – (covariance parameter for intercept of model with more elaborate vocal variable added / covariance parameter for intercept of model only including volubility); RVC = reciprocal vocal contingency (Harbison et al., 2018).

The communicative use composite accounted for a large incremental effect size regardless of the volubility measure that was controlled (pseudo R^2 change = .50 and .49) and was significant after controlling for total vocalizations from the communication samples, t(85.43) = 9.29, p < .001, and after controlling for the number of child speech-related vocalizations from the day-long audio recordings, t(83.88) = 9.39, p < .001. Neither of the volubility variables showed incremental validity in predicting later expressive language when the communicative use composite was controlled.

When the complexity composite from the conventional communication samples was used as the vocal complexity measure, it accounted for a large incremental effect size regardless of the volubility measure that was controlled (pseudo R^2 change = .61 and .60) and was statistically significant after controlling for total vocalizations from the communication samples, t(84.71) = 7.74, p < .001, and after controlling for number of child speech-related vocalizations, t(85.03) = 7.42, p < .001. Neither of the volubility variables showed incremental validity in predicting later expressive language when the complexity composite was controlled.

When the ACPU-C+V was used as the measure of complexity, it only showed incremental validity relative to the conventionally-coded volubility variable, $t(86.67) = 2.69$, $p < .01$. Total vocalizations measured conventionally was a significant predictor of end point expressive language, $t(87.07) = 3.93$, $p < .001$. ACPU-C+V did not exhibit incremental validity relative to the number of child speech-related vocalizations from the day-long audio recordings, $t(87.01) = 1.84$, $p = .07$. The number of child speech-related vocalizations was a significant predictor of end point expressive language, $t(84.41) = 3.69$, $p < .001$.

RVC incrementally predicted later expressive language only when total vocalizations was the measure of volubility that was controlled, $t(86.89) = 2.03$, $p < .05$. RVC did not predict later expressive language when the number of child speech-related vocalizations from day-long audio recordings were controlled ($p = .43$). The latter is noteworthy because the automated measure of volubility is the base rate of child vocalizations in the same day-long sample from which the RVC is derived.

Evaluating Incremental Validity of RVC Relative to Communicative Use and Complexity

To evaluate the incremental validity of the dyadic three-event variable, RVC, relative to the single actor, single event variables of communicative use and complexity, we added RVC to a model with only the communicative use composite from the communication samples or each of the two complexity variables (i.e., complexity composite from the communication samples or ACPU-C+V from the day-long audio recordings) to predict end point expressive language. See Table 10 for results. When either the communicative use or the complexity composite from the communication samples were statistically controlled, RVC no longer predicted end point expressive language, $p = .27$ or $p = .74$, respectively. In contrast, when ACPU-C+V, an automated measure of complexity, was controlled, RVC was a significant predictor of end point expressive language, $t(87.06) = 2.13$, $p =.04$; ACPU-C+V was a significant predictor as well, $t(86.64) = 2.08$, $p = .04$.

Table 10

Coefficient (Standard Errors) and Significance for RVC, a Dyadic Three-Event Variable, Predicting End Point Expressive Language After Controlling for Single Actor / Single Event Variables

Single Actor / Single Event Variable		RVC - A Dyadic Three-Event Variable	
Variable	Coefficient	Coefficient	Pseudo R^2 change for RVC
Communicative composite (conventional)	0.87 (0.09)***	0.72 (0.64)	.04
Complexity composite (conventional)	0.51 (0.06)***	0.24 (0.74)	.00
ACPU-C+V	0.21 (0.10)*	1.93 (0.93)*	.06

Note. * = $p < .05$; *** = $p < .001$; ACPU-C+V = Average count per utterance – consonants + vowels score (Woynaroski et al., 2017; Xu, Richards, & Gilkerson, 2014); Pseudo R^2 change = 1 – (covariance parameter for intercept of model with more elaborate vocal variable added / covariance parameter for intercept of model only including the less elaborate vocal variable); RVC = reciprocal vocal contingency (Harbison et al., 2018).

Evaluating Incremental Validity of Conventionally-Coded Measures Relative to Automated Measures of the Same Construct

To evaluate the incremental validity of conventionally-coded measures to automated measures of volubility and complexity, we added both measures in the same model to predict end point expressive language. See Table 11 for results. When the automated variable of the number of child speech-related vocalizations was added to the model with the conventionally-coded total vocalizations variable, total vocalizations and number of child speech-related vocalizations were significant predictors of end point expressive language, $t(88.21) = 2.66$, $p < .01$, and $t(85.11) = 3.07$, $p < .01$, respectively. When the automated complexity variable of ACPU-C+V was added to the model with the conventionally-coded complexity composite, only the complexity composite was a significant predictor of end point expressive language, $t(85.33) = 8.26$, $p < .001$. ACPU-C+V was not ($t(87.23) = 0.88$, $p = .38$). Adding the complexity composite to the model yielded a large effect size (pseudo R^2 change = .54).

Table 11

Coefficient (Standard Errors) and Significance for the Conventionally-Coded Vocal Variable Predicting End Point Expressive Language After Controlling for the Automated Vocal Variable

Automated Vocal Variable		Conventionally-Coded Vocal Variable		
Variable	Coefficient	Variable	Coefficient	Pseudo R^2 change
Number of child speech-related vocalizations	$2.2x10^4$ ($7.1x10^5$)**	Number of total vocalizations	0.19 (0.07)**	.10
ACPU-C+V	0.07 (0.08)	Complexity composite	0.50 (0.06)***	.54

Note. ** = p < .01; *** = p < .001; ACPU-C+V = average count per utterance – consonants + vowels (Woynaroski et al., 2017; Xu, Richards, & Gilkerson, 2014); Pseudo R^2 change = 1 – (covariance parameter for intercept of model with conventionally-coded vocal variable added / covariance parameter for intercept of model only including the automated vocal variable).

CHAPTER 4

DISCUSSION

Summary of Relative Validity of Vocal Variables

Because there is not yet consensus on how to compare the relative validity of competing variables, we present the results in multiple ways to aid readers with differing preferences. Nine vocal variables were evaluated for convergent validity, divergent validity, and sensitivity to change. See Table 12 for a substantive summary of the results.

Table 12

Summary of Evidence of Convergent Validity, Divergent Validity, and Sensitivity to Change

Construct	Vocal Variable	Evidence of Convergent Validity (RQ1)	Evidence of Divergent Validity (RQ2)	Evidence of Sensitivity to Change (RQ3)
Volubility	Number of total vocalizations	Yes	No	Yes$^+$
	Number of child speech-related vocalizations	Yes	Yes	Mixed
Communicative Use	Number of communication acts that include a vocalization	Yes$^+$	Yes	Yes$^+$
	Proportion of vocalizations that are communicative	Yes$^+$	Yes	Yes$^+$
Complexity	DKCC	Yes$^+$	Yes	Yes$^+$
	Proportion of vocalizations with a canonical syllable	Yes$^+$	Yes	Yes$^+$
	ACPU-C+V	Yes	Yes	Yes$^+$
	IVD score	No	Yes	No
Vocal Reciprocity	RVC	Yes	Yes	Yes

Note. $^+$ = large effect size, defined as pseudo $R^2 \geq .25$ (for convergent validity) or Cohen's $d \geq .8$ (for sensitivity to change); ACPU-C+V = average count per utterance – consonants + vowels (Woynaroski et al., 2017; Xu, Richards, & Gilkerson, 2014); DKCC = diversity of key consonants used in communication acts (Wetherby, Watt, Morgan, & Shumway, 2007; Woynaroski et al., 2017); IVD = infraphonological vocal development (Oller et al., 2010); RVC = reciprocal vocal contingency (Harbison et al., 2018).

If we use the presence or absence of significant results as the criterion for assigning validity and summarizing across these three purposes, six variables presented with consistent positive evidence for all three purposes (i.e., number of communication acts that include a vocalization, proportion of vocalizations that are communicative, DKCC, proportion of vocalizations with a canonical syllable, ACPU-C+V, and RVC). However, we do not recommend using the number of validity tests as the basis for selecting vocal variables. Doing so ignores the relative effect size of the associations or change and may result in ignoring variables that may be most valid for particular purposes. Thus, effect-size informed, purpose-specific decisions are more likely to be useful. Using above-threshold effect size as a basis for selecting variables, four variables showed evidence of validity to predict expressive language and show sensitivity to change (i.e., number of communication acts that include a vocalization, proportion of vocalizations that are communicative, DKCC, and proportion of vocalizations with a canonical syllable). Note that these four variables are all conventionally-coded, two measure communicative use and two measure complexity. DKCC had the largest effect size for convergent validity and for sensitivity to change.

If we instead use criteria of the strongest evidence for construct validity (i.e., convergent and divergent validity together), then the variables that are most supported are the two measures of communicative use (i.e., number of communicative acts that include a vocalization and proportion of vocalizations that are communicative) and the two conventionally-coded measures of complexity (i.e., DKCC and proportion of vocalizations with a canonical syllable). If we use the criterion of which variables show a large effect size for sensitivity to change, then six variables are most supported (i.e., number of total vocalizations, number of communication acts that include a vocalization, proportion of vocalizations that are communicative, DKCC, proportion of vocalizations with a canonical syllable, and ACPU-C+V).

If we choose to identify which variables show the strongest evidence of convergent validity (i.e., prediction of later expressive language), our most rigorous evidence is provided by tests of incremental validity. See Table 13 for a substantive summary of the incremental validity results.

Table 13

Summary of Evidence of Incremental Validity for Predicting Later Expressive Language

Comparison	Less Elaborate/Less Costly Variable	More Elaborate/More Costly Variable
More elaborate / more costly versus volubility	[ns]Volubility (conventional or automated)	*Communicative composite (conventional)
	[ns]Volubility (conventional or automated)	*Complexity composite (conventional)
	*Number of total vocalizations (conventional)	*ACPU-C+V (automated)
	*Number of child speech-related vocalizations (automated)	[ns]ACPU-C+V (automated)
	*Number of total vocalizations (conventional)	*RVC (automated)
	*Number of child speech-related vocalizations (automated)	[ns]RVC (automated)
RVC versus communicative use or complexity	*Communicative composite (conventional)	[ns]RVC (automated)
	*Complexity composite (conventional)	[ns]RVC (automated)
	*ACPU-C+V (automated)	*RVC (automated)
Automated versus conventionally-coded	*Number of child speech-related vocalizations (automated)	*Number of total vocalizations (conventional)
	[ns]ACPU-C+V (automated)	*Complexity composite (conventional)

Note. * = significant incremental validity; [ns] = non-significant incremental validity; ACPU-C+V = average count per utterance – consonants + vowels (Woynaroski et al., 2017; Xu, Richards, & Gilkerson, 2014); RVC = reciprocal vocal contingency (Harbison et al., 2018).

For the purpose of predicting expressive language, measuring complexity or communicative use of vocalizations has incremental validity after controlling for volubility. To estimate vocal complexity, it is likely worth the effort to conventionally code communication samples. The incremental validity of the automated measure of vocal complexity (i.e., ACPU-C+V) was nonsignificant after controlling for the automated measure of volubility.

Incremental validity comparisons of automated versus conventionally-coded variables allowed us to test whether it was worth the cost of conventionally-coded communication samples to measure similar constructs. We found evidence of incremental validity of the automated volubility variable when compared to the conventional measure of volubility. However, the value of investing in equipment for an automated measure of volubility ought to be considered in the context of the incremental validity of predicting expressive language controlling for other vocal variables. As mentioned earlier, there is weaker or mixed evidence to consider including measures of volubility in a battery of vocal variables when evaluated from the perspective of overall validity and purpose-specific validity as compared to conventional measures of complexity and communicative use.

In contrast to the automated versus conventional volubility comparison results, only the conventionally-coded complexity measure accounted for unique variance in predicting end point expressive language skills when both the automated (i.e., ACPU-C+V) and conventionally-coded variables were in the model. These results provide additional evidence in support of conventionally-coded vocal complexity measures, despite the resources required.

Another important finding appeared when considering the full set of results. The findings do not support the construct validity of IVD score for measuring vocal development of young children with ASD who are in the early stages of word learning. IVD score did not exhibit evidence of convergent validity or sensitivity to change. These results bring into question the validity of the IVD score for assessing vocal development in young children with ASD. They conflict with evidence supporting the convergent validity of IVD score, as described in more detail below (Woynaroski et al., 2017).

Similarly, these data do not support using the RVC as an incrementally-valid predictor of expressive language in children with ASD and conflict with some earlier findings. As indicated in the introduction, RVC was designed to be predict convergent validity variables such as expressive language after controlling for the chance sequential occurrence of adult and child vocalizations (which is computed from their base rates). In this study, RVC was not an incrementally-valid predictor of expressive language after controlling for the automated measure of child volubility. This finding conflicts with past evidence that RVC was related to expressive communication in a smaller sample of children with ASD (*n* = 21; Harbison et al., 2018). Reasons to consider current study findings as more informative of the validity of RVC than Harbison et al. (2018) include the larger sample size, the more informative research design, and a multi-measure approach to estimating expressive language in the current study.

The Current Study Findings Relative to the Extant Literature

The pattern of results highlights the value of considering the communicative use and complexity of vocalizations when evaluating vocal development in young children with ASD. The value observed in considering communicative use and/or complexity aligns with the social feedback theory, the speech attunement framework, and the transactional theory of spoken language development. However, there are also child-driven theories that support the selection of communicative use and complexity vocal variables as putative predictors of expressive language. Thus, the current study does not support one theory over another in a definitive way. As indicated in the introduction, there is much converging evidence across studies to support the validity of measuring complexity and communicative use of vocalizations. In this section we identify specific findings within the extant literature for which the current findings are replications. See Table 14 for a summary.

Table 14

Replicated Findings Identified in Extant Literature and Current Study

Vocal Variable	Convergent Validity	Divergent Validity	Sensitivity to Change
Number of total vocalizations	Plumb (2008)	N/A	Brian et al. (2017)
Number of child speech-related vocalizations			Dykstra et al. (2013)
Number of communication acts that include a vocalization	Plumb (2008)		
Proportion of vocalizations that are communicative			
DKCC	Wetherby et al. (2007) Woynaroski (2014) & Woynaroski et al. (2017) Yoder et al. (2015)		Woynaroski et al. (2016)
Proportion of vocalizations with a canonical syllable			
ACPU-C+V	Woynaroski et al. (2017)		
IVD score	N/A		N/A
RVC	Harbison et al. (2018)	Harbison et al. (2018)	

Note. N/A = not applicable because finding not identified in current study; ACPU-C+V = average count per utterance – consonants + vowels (Woynaroski et al., 2017; Xu, Richards, & Gilkerson, 2014); DKCC = diversity of key consonants used in communication acts (Wetherby, Watt, Morgan, & Shumway, 2007; Woynaroski et al., 2017); IVD = infraphonological vocal development (Oller et al., 2010); RVC = reciprocal vocal contingency (Harbison et al., 2018).

Convergent validity based on significant correlations with current or later expressive language has been reported for the number of total vocalizations (Plumb, 2008), the number communication acts that include a vocalization (Plumb, 2008), DKCC (Wetherby et al., 2007; Woynaroski, 2014; Woynaroski et al., 2017; Yoder et al., 2015), ACPU-C+V (Woynaroski et al., 2017), and RVC (Harbison et al., 2018). DKCC has previously exhibited incremental validity as well (Yoder et al., 2015). Thus, these findings from the current study are replications, decreasing the likelihood that they are sample specific results and increasing the likelihood for future replications. In contrast, Plumb and Wetherby (2013) did not identify a significant relation between proportion of vocalizations that are communicative and expressive language as measured by the verbal development quotient on the MSEL. The current study may have identified this previously unidentified relation in part due to the relatively large sample size and use of growth curve modeling to generate a better estimate of end point expressive language than an observed measure at a single time point (Singer & Willett, 2003). Additionally, the current study quantified expressive language using multiple measures, in contrast to the sole use of the MSEL in Plumb and Wetherby (2013).

No known extant studies report convergent validity for assessing vocal development in young children with ASD for the proportion of vocalizations that include a canonical syllable. However, Williams (2013) evaluated a variable similar to the proportion of vocalizations that include a canonical syllable in testing the correlation between the percent of syllables that are canonical and the MSEL Early Language Composite and the MSEL Expressive Language Composite for 15 infant siblings of children with ASD who were 6 months old and at high-risk for ASD. This vocal variable is similar to, but not synonymous with, proportion of vocalizations that include a consonant. The results were non-significant for the Early Language and Expressive Language composites, r = -.21 and r = .21, respectively. The fact that most infant siblings will not have ASD and the relatively younger and smaller sample as well as the use of different language measures relative to the current study may explain the incongruent findings.

The results of the current study also differ from Woynaroski et al. (2017) who reported that IVD score exhibited predictive validity for spoken vocabulary. In Woynaroski et al. (2017) IVD score and ACPU-C+V, another automatic putative measure of vocal complexity, correlated sufficiently (r = .45; p = .023) to aggregate; however, in the current study they did not. Possible reasons for differences between the current study and Woynaroski et al. (2017) include differences in the number of day-long audio recordings per child, the interval between the IVD score and expressive language measurement, and the characteristics of the sample. In Woynaroski et al. (2017), two day-long audio recordings were used to estimate IVD score. Only one day-long audio sample per child per time period was available for the current study. In addition, Woynaroski et al. (2017) tested the correlation between the vocal variables and expressive vocabulary 4 months later. We used multi-level modeling to test whether the vocal variables predicted the best estimate of expressive language 12 months later. Participants in Woynaroski et al. (2017) were all preverbal (i.e., no more than 20 spoken words at study initiation), whereas the participants in the current study exhibited more varied expressive language skills at study initiation.

Comparisons between the extant literature and current findings for divergent validity are very limited. The only known study to assess the divergent validity of a vocal variable for assessing vocal development in young children with ASD is Harbison et al. (2018), which evaluated divergent validity of RVC. Although Harbison et al. (2018) used different divergent validity nodes (i.e., chronological age, intellectual quotient, and parents' formal education level) than the current study, the current study and Harbison et al. (2018) provide positive evidence of divergent validity for RVC assessing vocal development in young children with ASD. Unfortunately, the lack of incremental validity of the RVC in the current study weakens the support for the RVC as a measure of reciprocal vocal interaction.

Evidence for sensitivity to change of vocal variables in children with ASD change can be drawn from studies examining change in vocal variables over time, even when the study does not specifically identify the analysis as one of sensitivity to change. Brian, Smith, Zwaigenbaum and Bryson (2017) conducted a randomized control trial on the efficacy of *Social ABCs*, a parent-mediated intervention. They reported increased child vocal initiations from the beginning to the end of the 12-week intervention period ($z = 4.206$, $p < .001$). Using an automated measure of volubility, Dykstra et al. (2013) reported that the rate of speech-related child vocalizations per minute from day-long LENA recordings increased from the beginning to the end of the school year for a sample of 40 children with ASD (mean age = 3.95 years). This rate measure is analogous to our automated measure of volubility. For a measure of vocal complexity, Woynaroski et al. (2016) reported significant simple linear growth in DKCC for 87 initially preverbal children with ASD across 16 months in a longitudinal correlational study. Thus, the evidence for sensitivity to change for DKCC in the current sample replicates this finding.

Limitations

Five 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, uses, or populations. The current investigation selected vocal variables for the purpose of measuring vocal development in young children with ASD in the early stages of language learning. 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 the number of variables was reduced to partially address family-wise error due to multiple significance tests without alpha adjustment, there are still many significance tests per research question. Replication 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 (e.g., Harbison et al., 2018; McDaniel et al., 2018; Plumb & Wetherby, 2013; Wetherby et al., 2007; Woynaroski et al., 2017; Yoder et al., 2015). It is unlikely that these associations have been found to be significant due to unadjusted multiple significance testing. Third, only one day-long LENA recording was collected per participant per time period. Thus, the degree of stability across days could not be assessed and the use of a single sample may not have been sufficiently stable to optimize the automated variables. Results may have differed if two or more day-long samples per participant per time point were utilized, particularly for RVC. RVC has been shown to be stable across two day-long samples in children with ASD (Harbison et al., 2018). However, past work has shown that IVD is stable with only one day-long sample

(Woynaroski et al., 2017; Yoder et al., 2013). Thus, the absence of validity for IVD score is not solely due to the use of a single day-long sample. Fourth, clear divisions among variables that measure communicative use versus complexity were not possible in every case. For example, DKCC considers communicative use as well as complexity. Even so, there was incremental validity of complexity and communicative use after controlling for the other. Fifth, 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 the treatments caused the change in the vocal variables. The one exception is that it appears that the high intensity treatment caused more change in the number of child speech-related vocalizations than did the low intensity treatment. However, the lack of theoretical rationale for intensity or style of treatment affecting volubility, only when measured using the automated method and the large number of significance tests, lead us to conclude this finding was likely sample specific.

Strengths

Six 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 the literature on children with ASD and vocal measures in any population. Second, this study includes conventionally-coded and automated vocal variables for the same participants, which enables direct comparisons of two ways to derive volubility and complexity. Third, 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 & Willett, 2003). Fourth, 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. This current study design permitted addressing predictive validity, one important purpose for which vocal variables are often needed. Fifth, this study includes a relatively large sample size for this population, which increased the power to detect effects and permitted the use of multi-level models with the necessary number of predictors to address the research questions. Sixth, multiple ways to compare the validity of multiple vocal variables were presented to meet the different needs of readers with different perspectives regarding how to select among the many vocal variables.

Implications

Consistent with the rationale for conducting this study, 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 communicative use and complexity when assessing vocal development in young children with ASD, particularly when derived by human coding of communication samples. One can use these findings when selecting study variables for different purposes for which assessing and targeting vocal development may be useful. These potential purposes include increasing the effectiveness of communication intervention, identifying early response to intervention, explaining why intervention is effective, and for whom intervention is effective in initially preverbal children with ASD. For example, when selecting variables that might mediate treatment effects on expressive language, the findings suggest that using variables of communicative use and complexity may maximize the probability of detecting the putative mediated effect of early language interaction on expressive language through midpoint vocal development. As another example, one might consider using variables that demonstrate sensitivity to change to assess progress in vocal complexity (e.g., proportion of vocalizations that are communicative or DKCC) over a vocal complexity variable that did not exhibit sensitivity to change (e.g., IVD score).

Future Directions

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 skills. Relatedly, because construct validity is judged based on a network of nodes (Cronbach & Meehl, 1955), additional nodes for convergent and divergent validity may be explored to increase the confidence in the current results. For instance, correlations with receptive language may be considered for convergent validity nodes. It may not be intuitive to predict that vocal variables may predict receptive language, but past work has shown that DKCC was predicted by receptive language (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. Given the mixed evidence for the construct validity of RVC for young children with ASD in the early stages of language learning, additional work to optimize this measure is warranted. For example, at least two day-long samples might be needed to improve the incremental validity of the RVC over volubility. Future studies could investigate whether there is a particular communication or language level for which RVC is most valid as well as how to minimize error from the sampling procedures (e.g., using multiple day-long samples per time point).

Although general categories of cost were determined for the current study, detailed cost analyses were not possible. Future studies should also consider the specific cost of variables to inform variable selection and planning of later investigations. Relatedly, 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 for the broader scientific community to measure vocal development within and across young children with ASD. Key findings include strong evidence of construct validity and incremental validity for predicting expressive language using conventional methods to measure complexity and communicative use of vocalizations in young children with ASD. These results support the use of conventional measures of complexity and communicative use of vocalizations in future studies of language intervention in children with ASD.

Appendix A

Intercorrelations Between Vocal Variables of Communicative Use and Complexity

Intercorrelations Between Communicative Use Vocal Variables at Time 1

	1	2	3
Number of CAs that include a vocalization	1		
Number of CAs that include a CS	.985**	1	
Proportion of vocalizations that are communicative	.699**	.673**	1

Note. ** = p < .01; CA = communication act; CS = canonical syllable.

Intercorrelations Between Communicative Use Vocal Variables at Time 3

	1	2	3
Number of CAs that include a vocalization	1		
Number of CAs that include a CS	.996**	1	
Proportion of vocalizations that are communicative	.789**	.803**	1

Note. ** = p < .01; CA = communication act; CS = canonical syllable.

Intercorrelations Between Conventionally-Coded Vocal Variables for Complexity at Time 1

	1	2	3	4	5
1. Consonant inventory	1				
2. DKCC	.853**	1			
3. Proportion of CAs with a CS	.682**	.633**	1		
4. Proportion of vocalizations with a CS	.828**	.789**	.713**	1	
5. Number of vocalizations with a CS	.847**	.857**	.618**	.769**	1

Note. ** = p < .01; CA = communication act; CS = canonical syllable; DKCC = diversity of key consonants used in communication acts (Wetherby, Watt, Morgan, & Shumway, 2007; Woynaroski et al., 2017).

Intercorrelations Between Conventionally-Coded Vocal Variables for Complexity at Time 3

	1	2	3	4	5
1. Consonant inventory	1				
2. DKCC	.878**	1			
3. Proportion of CAs with a CS	.714**	.665**	1		
4. Proportion of vocalizations with a CS	.818**	.872**	.757**	1	
5. Number of vocalizations with a CS	.752**	.750**	.512**	.675**	1

Note. ** = p <.01; CA = communication act; CS = canonical syllable; DKCC = diversity of key consonants used in communication acts (Wetherby, Watt, Morgan, & Shumway, 2007; Woynaroski et al., 2017).

Intercorrelations Between Automated Vocal Variables for Complexity at Time 1

	1	2	3
1. ACPU-C	1		
2. ACPU-V	.824**	1	
3. IVD score	.123	.361**	1

Note. ** = p < .01; ACPU-C = average count per utterance – consonants (Xu, Richards, & Gilkerson, 2014); ACPU-V = average count per utterance – vowels (Xu, Richards, & Gilkerson, 2014); IVD = infraphonological vocal development (Oller et al., 2010).

Intercorrelations Between Automated Vocal Variables for Complexity at Time 3

	1	2	3
1. ACPU-C	1		
2. ACPU-V	.910**	1	
3. IVD score	-.056	.007	1

Note. ** = p < .01; ACPU-C = average count per utterance – consonants (Xu, Richards, & Gilkerson, 2014); ACPU-V = average count per utterance – vowels (Xu, Richards, & Gilkerson, 2014); IVD = infraphonological vocal development (Oller et al., 2010).

Appendix B

Growth Curve Model Results for Convergent Validity Analyses

Estimates of Fixed Effects for Empty Model

Parameter	Estimate	*SE*	*df*	*t*	*p*	Lower CI	Upper CI
Intercept	-0.66	0.07	86.48	-9.80	<.001	-0.79	-0.53

Covariance Parameters for Empty Model

Parameter	Estimate	*SE*	Wald Z	*p*	Lower CI	Upper CI
Residual	0.58	0.06	9.10	<.001	0.47	0.72
Intercept	0.19	0.06	3.01	<.01	0.10	0.36

Estimates of Fixed Effects for Random Intercept Fixed Slope Model

Parameter	Estimate	*SE*	*df*	*t*	*p*	Lower CI	Upper CI
Intercept	0.01	0.07	132.26	0.11	.92	-0.14	0.15
Time	0.11	0.01	167.98	20.49	<.001	0.10	0.12

Covariance Parameters for Random Intercept Fixed Slope Model

Parameter	Estimate	*SE*	Wald Z	*p*	Lower CI	Upper CI
Residual	0.16	0.02	9.10	<.001	0.13	0.20
Intercept	0.31	0.06	5.57	<.001	0.22	0.45

Estimates of Fixed Effects for Model with Total Vocalizations as Predictor Variable

Parameter	Estimate	*SE*	*df*	*t*	*p*	Lower CI	Upper CI
Intercept	0.21	0.09	119.83	2.36	.02	0.03	0.38
Time	0.11	0.01	167.02	20.51	<.001	0.10	0.12
Total vocalizations	0.26	0.07	88.57	3.74	<.001	0.12	0.40

Covariance Parameters for Model with Total Vocalizations as Predictor Variable

Parameter	Estimate	*SE*	Wald Z	*p*	Lower CI	Upper CI
Residual	0.16	0.02	9.08	<.001	0.13	0.20
Intercept	0.26	0.05	5.36	<.001	0.18	0.38

Estimates of Fixed Effects for Model with Number of Child Speech-Related Vocalizations as Predictor Variable

Parameter	Estimate	*SE*	*df*	*t*	*p*	Lower CI	Upper CI
Intercept	-0.47	0.13	95.74	-3.57	.001	-0.72	-0.21
Time	0.11	0.01	163.29	20.23	<.001	0.10	0.12
Number of child speech-related vocalizations	2.9×10^4	6.8×10^5	83.89	4.25	<.001	1.5×10^4	4.3×10^4

Covariance Parameters for Model with Number of Child Speech-Related Vocalizations as Predictor Variable

Parameter	Estimate	*SE*	Wald Z	*p*	Lower CI	Upper CI
Residual	0.16	0.02	8.98	<.001	0.13	0.20
Intercept	0.25	0.05	5.29	<.001	0.18	0.37

Estimates of Fixed Effects for Model with Number of Communication Acts that Include a Vocalization as Predictor Variable

Parameter	Estimate	*SE*	*df*	*t*	*p*	Lower CI	Upper CI
Intercept	0.63	0.09	112.58	6.82	<.001	0.45	0.81
Time	0.11	0.01	168.00	20.59	<.001	0.10	0.12
Number of communication acts that include a vocalization	0.86	0.10	86.16	8.64	<.001	0.66	1.05

Covariance Parameters for Model with Number of Communication Acts that Include a Vocalization as Predictor Variable

Parameter	Estimate	*SE*	Wald Z	*p*	Lower CI	Upper CI
Residual	0.16	0.02	9.09	<.001	0.13	0.20
Intercept	0.14	0.03	4.62	<.001	0.09	0.22

Estimates of Fixed Effects for Model with Proportion of Vocalizations that are Communicative as Predictor Variable

Parameter	Estimate	*SE*	*df*	*t*	*p*	Lower CI	Upper CI
Intercept	0.56	0.08	118.55	6.84	<.001	0.40	0.72
Time	0.11	0.01	168.63	20.54	<.001	0.10	0.12
Proportion of vocalizations that are communicative	0.69	0.07	85.55	9.35	<.001	0.55	0.84

Covariance Parameters for Model with Proportion of Vocalizations that are Communicative as Predictor Variable

Parameter	Estimate	*SE*	Wald Z	*p*	Lower CI	Upper CI
Residual	0.16	0.02	9.11	<.001	0.13	0.20
Intercept	0.13	0.03	4.49	<.001	0.08	0.20

Estimates of Fixed Effects for Model with Diversity of Key Consonants Used in Communication Acts as Predictor Variable

Parameter	Estimate	*SE*	*df*	*t*	*p*	Lower CI	Upper CI
Intercept	0.67	0.09	115.54	7.43	<.001	0.49	0.85
Time	0.11	0.01	166.38	20.62	<.001	0.10	0.12
DKCC	0.59	0.06	85.86	9.41	<.001	0.47	0.72

Note. DKCC = diversity of key consonants used in communication acts (Wetherby, Watt, Morgan, & Shumway, 2007; Woynaroski et al., 2017).

Covariance Parameters for Model with Diversity of Key Consonants Used in Communication Acts as Predictor Variable

Parameter	Estimate	*SE*	Wald Z	*p*	Lower CI	Upper CI
Residual	0.17	0.02	9.05	<.001	0.13	0.21
Intercept	0.13	0.03	4.35	<.001	0.08	0.20

Estimates of Fixed Effects for Model with Proportion of Vocalizations with a Canonical Syllable as Predictor Variable

Parameter	Estimate	*SE*	*df*	*t*	*p*	Lower CI	Upper CI
Intercept	0.38	0.08	122.95	4.76	<.001	0.22	0.54
Time	0.11	0.01	167.90	20.55	<.001	0.10	0.12
Proportion of vocalizations with a canonical syllable	0.38	0.05	86.20	7.26	<.001	0.28	0.48

Covariance Parameters for Model with Proportion of Vocalizations with a Canonical Syllable as Predictor Variable

Parameter	Estimate	*SE*	Wald Z	*p*	Lower CI	Upper CI
Residual	0.16	0.02	9.10	<.001	0.13	0.20
Intercept	0.18	0.04	4.90	<.001	0.12	0.26

Estimates of Fixed Effects for Model with ACPU-C+V as Predictor Variable

Parameter	Estimate	*SE*	*df*	*t*	*p*	Lower CI	Upper CI
Intercept	0.19	0.10	103.79	1.93	.06	-0.01	0.40
Time	0.11	0.01	163.23	20.20	<.001	0.10	0.12
ACPU-C+V	0.27	0.10	86.50	2.71	<.01	0.07	0.46

Note. ACPU-C+V = average count per utterance – consonants + vowels (Woynaroski et al., 2017; Xu, Richards, & Gilkerson, 2014).

Covariance Parameters for Model with ACPU-C+V as Predictor Variable

Parameter	Estimate	*SE*	Wald Z	*p*	Lower CI	Upper CI
Residual	0.16	0.02	8.98	<.001	0.13	0.20
Intercept	0.29	0.05	5.43	<.001	0.20	0.42

Note. ACPU-C+V = average count per utterance – consonants + vowels (Woynaroski et al., 2017; Xu, Richards, & Gilkerson, 2014).

Estimates of Fixed Effects for Model with Infraphonological Vocal Development Score as Predictor Variable

Parameter	Estimate	*SE*	*df*	*t*	*p*	Lower CI	Upper CI
Intercept	0.04	0.14	94.11	0.29	.77	-0.24	0.32
Time	0.11	0.01	163.07	20.22	<.001	0.10	0.12
Infraphonological vocal development score	0.00	0.01	84.66	-0.32	.75	-0.02	0.01

Covariance Parameters for Model with Infraphonological Vocal Development Score as Predictor Variable

Parameter	Estimate	*SE*	Wald Z	*p*	Lower CI	Upper CI
Residual	0.16	0.02	8.98	<.001	0.13	0.20
Intercept	0.32	0.06	5.51	<.001	0.22	0.46

Estimates of Fixed Effects for Model with Reciprocal Vocal Contingency as Predictor Variable

Parameter	Estimate	*SE*	*df*	*t*	*p*	Lower CI	Upper CI
Intercept	-0.51	0.21	90.93	-2.48	.02	-0.92	-0.10
Time	0.11	0.01	162.83	20.19	<.001	0.10	0.12
Reciprocal vocal contingency	2.45	0.92	85.67	2.67	<.01	0.62	4.27

Covariance Parameters for Model with Reciprocal Vocal Contingency as Predictor Variable

Parameter	Estimate	*SE*	Wald Z	*p*	Lower CI	Upper CI
Residual	0.16	0.02	8.97	<.001	0.13	0.20
Intercept	0.29	0.05	5.41	<.001	0.20	0.42

Appendix C

Growth Curve Model Results for Divergent Validity Analyses

Estimates of Fixed Effects for Empty Model

Parameter	Estimate	*SE*	*df*	*t*	*p*	Lower CI	Upper CI
Intercept	-0.67	0.07	86.66	-10.33	<.001	-0.80	-0.54

Covariance Parameters for Empty Model

Parameter	Estimate	*SE*	Wald Z	*p*	Lower CI	Upper CI
Residual	0.58	0.06	9.04	<.001	0.47	0.73
Intercept	0.16	0.06	2.72	<.01	0.08	0.33

Estimates of Fixed Effects for Random Intercept Fixed Slope Model

Parameter	Estimate	*SE*	*df*	*t*	*p*	Lower CI	Upper CI
Intercept	0.04	0.07	124.17	0.51	.61	-0.10	0.17
Time	0.11	0.00	165.19	25.38	<.001	0.11	0.12

Covariance Parameters for Random Intercept Fixed Slope Model

Parameter	Estimate	*SE*	Wald Z	*p*	Lower CI	Upper CI
Residual	0.12	0.01	9.02	<.001	0.10	0.15
Intercept	0.30	0.05	5.79	<.001	0.21	0.42

Estimates of Fixed Effects for Random Intercept Random Slope Model

Parameter	Estimate	*SE*	*df*	*t*	*p*	Lower CI	Upper CI
Intercept	0.03	0.09	86.99	0.39	.70	-0.14	0.20
Time	0.11	0.01	80.33	21.35	.00	0.10	0.12

Covariance Parameters for Random Intercept Random Slope Model

Parameter		Estimate	*SE*	Wald Z	*p*	Lower CI	Upper CI
Residual		0.07	0.01	6.29	<.001	0.05	0.10
Intercept + Time	UN (1,1)	0.55	0.09	5.81	<.001	0.39	0.77
	UN (2,1)	0.02	0.01	4.31	<.001	0.01	0.03
	UN (2,2)	0.001	4.0x10^{4}	3.24	.001	7.1x10^{4}	0.002

Estimates of Fixed Effects for Model with Total Vocalizations as Predictor Variable

Parameter	Estimate	*SE*	*df*	*t*	*p*	Lower CI	Upper CI
Intercept	0.11	0.09	106.72	1.23	.22	-0.07	0.30
Time	0.11	0.01	79.20	21.33	<.001	0.10	0.12
Total vocalizations	0.13	0.06	86.42	2.16	.03	0.01	0.25

Covariance Parameters for Model with Total Vocalizations as Predictor Variable

Parameter		Estimate	*SE*	Wald Z	*p*	Lower CI	Upper CI
Residual		0.07	0.01	6.25	<.001	0.05	0.10
Intercept + Time	UN (1,1)	0.51	0.09	5.66	<.001	0.36	0.72
	UN (2,1)	0.02	0.01	4.17	<.001	0.01	0.03
	UN (2,2)	0.001	$4.0\text{x}10^{4}$	3.16	<.01	$6.7\text{x}10^{4}$	0.002

Estimates of Fixed Effects for Model with Number of Child Speech-Related Vocalizations as Predictor Variable

Parameter	Estimate	*SE*	*df*	*t*	*p*	Lower CI	Upper CI
Intercept	-0.10	0.13	125.25	-0.79	.43	-0.35	0.15
Time	0.11	0.01	77.47	20.82	<.001	0.10	0.12
Number of child speech-related vocalizations	$6.7\text{x}10^{5}$	$5.8\text{x}10^{5}$	84.49	1.16	.25	$-4.8\text{x}10^{5}$	$1.8\text{x}10^{4}$

Covariance Parameters for Model with Number of Child Speech-Related Vocalizations as Predictor Variable

Parameter		Estimate	*SE*	Wald Z	*p*	Lower CI	Upper CI
Residual		0.07	0.01	6.17	<.001	0.05	0.10
Intercept + Time	UN (1,1)	0.54	0.10	5.69	<.001	0.38	0.76
	UN (2,1)	0.02	0.01	4.31	<.001	0.01	0.03
	UN (2,2)	0.001	$4.1\text{x}10^{4}$	3.15	<.01	$6.9\text{x}10^{4}$	0.002

Estimates of Fixed Effects for Model with Number of Communication Acts that Include a Vocalization as Predictor Variable

Parameter	Estimate	*SE*	*df*	*t*	*p*	Lower CI	Upper CI
Intercept	0.06	0.12	122.27	0.55	.58	-0.17	0.29
Time	0.11	0.01	79.30	21.33	<.001	0.10	0.12
Number of communication acts that include vocalization	0.06	0.11	86.09	0.54	.59	-0.16	0.28

Covariance Parameters for Model with Number of Communication Acts that Include a Vocalization as Predictor Variable

Parameter		Estimate	*SE*	Wald Z	*p*	Lower CI	Upper CI
Residual		0.07	0.01	6.25	<.001	0.05	0.10
Intercept + Time	UN (1,1)	0.54	0.09	5.77	<.001	0.39	0.76
	UN (2,1)	0.02	0.01	4.25	<.001	0.01	0.03
	UN (2,2)	0.001	$4.0\text{x}10^{4}$	3.16	<.01	$6.7\text{x}10^{4}$	0.002

Estimates of Fixed Effects for Model with Proportion of Vocalizations that are Communicative as Predictor Variable

Parameter	Estimate	*SE*	*df*	*t*	*p*	Lower CI	Upper CI
Intercept	-0.03	0.11	118.62	-0.26	.79	-0.24	0.19
Time	0.11	0.01	79.22	21.32	<.001	0.10	0.12
Proportion of vocalizations that are communicative	-0.06	0.09	86.85	-0.75	.46	-0.23	0.11

Covariance Parameters for Model with Proportion of Vocalizations that are Communicative as Predictor Variable

Parameter		Estimate	*SE*	Wald Z	*p*	Lower CI	Upper CI
Residual		0.07	0.01	6.25	<.001	0.05	0.10
Intercept + Time	UN (1,1)	0.55	0.10	5.77	<.001	0.39	0.77
	UN (2,1)	0.02	0.01	4.26	<.001	0.01	0.03
	UN (2,2)	0.001	4.0x10^{4}	3.17	<.01	6.7x10^{4}	0.002

Estimates of Fixed Effects for Model with Diversity of Key Consonants Used in Communication Acts as Predictor Variable

Parameter	Estimate	*SE*	*df*	*t*	*p*	Lower CI	Upper CI
Intercept	0.04	0.12	122.26	0.35	.73	-0.19	0.27
Time	0.11	0.01	79.31	21.33	<.001	0.10	0.12
DKCC	0.02	0.07	86.14	0.25	.81	-0.13	0.16

Note. DKCC = diversity of key consonants used in communication acts (Wetherby, Watt, Morgan, & Shumway, 2007; Woynaroski et al., 2017).

Covariance Parameters for Model with Diversity of Key Consonants Used in Communication Acts as Predictor Variable

Parameter		Estimate	*SE*	Wald Z	*p*	Lower CI	Upper CI
Residual		0.07	0.01	6.25	<.001	0.05	0.10
Intercept + Time	UN (1,1)	0.55	0.09	5.77	<.001	0.39	0.77
	UN (2,1)	0.02	0.01	4.25	<.001	0.01	0.03
	UN (2,2)	0.001	4.0x10^{4}	3.16	<.01	6.7x10^{4}	0.002

Estimates of Fixed Effects for Model with Proportion of Vocalizations with a Canonical Syllable as Predictor Variable

Parameter	Estimate	*SE*	*df*	*t*	*p*	Lower CI	Upper CI
Intercept	-0.01	0.10	113.38	-0.15	.89	-0.21	0.18
Time	0.11	0.01	79.27	21.32	<.001	0.10	0.12
Proportion of vocalizations with a canonical syllable	-0.04	0.05	86.14	-0.70	.49	-0.14	0.07

Covariance Parameters for Model with Proportion of Vocalizations with a Canonical Syllable as Predictor Variable

Parameter		Estimate	*SE*	Wald Z	*p*	Lower CI	Upper CI
Residual		0.07	0.01	6.25	<.001	0.05	0.10
Intercept + Time	UN (1,1)	0.54	0.09	5.77	<.001	0.39	0.76
	UN (2,1)	0.02	0.01	4.25	<.001	0.01	0.03
	UN (2,2)	0.001	4.0x10^{4}	3.16	<.01	6.7x10^{4}	0.002

Estimates of Fixed Effects for Model with ACPU-C+V as Predictor Variable

Parameter	Estimate	*SE*	*df*	*t*	*p*	Lower CI	Upper CI
Intercept	-0.06	0.10	112.57	-0.61	.54	-0.27	0.14
Time	0.11	0.01	77.33	20.81	<.001	0.10	0.12
ACPU-C+V	-0.10	0.08	84.85	-1.24	.22	-0.25	0.06

Note. ACPU-C+V = average count per utterance – consonants + vowels (Woynaroski et al., 2017; Xu, Richards, & Gilkerson, 2014).

Covariance Parameters for Model with ACPU-C+V as Predictor Variable

Parameter		Estimate	*SE*	Wald Z	*p*	Lower CI	Upper CI
Residual		0.07	0.01	6.17	<.001	0.05	0.10
Intercept + Time	UN (1,1)	0.53	0.09	5.65	<.001	0.37	0.75
	UN (2,1)	0.02	0.01	4.27	<.001	0.01	0.03
	UN (2,2)	0.001	4.0x10^{4}	3.15	<.01	6.9x10^{4}	0.002

Note. ACPU-C+V = average count per utterance – consonants + vowels (Woynaroski et al., 2017; Xu, Richards, & Gilkerson, 2014).

Estimates of Fixed Effects for Model with Infraphonological Vocal Development Score as Predictor Variable

Parameter	Estimate	*SE*	*df*	*t*	*p*	Lower CI	Upper CI
Intercept	-0.09	0.12	124.51	-0.69	.49	-0.33	0.16
Time	0.11	0.01	77.41	20.82	<.001	0.10	0.12
Infraphonological vocal development score	0.01	0.01	84.53	1.04	.30	-0.01	0.02

Covariance Parameters for Model with Infraphonological Vocal Development Score as Predictor Variable

Parameter		Estimate	*SE*	Wald Z	*p*	Lower CI	Upper CI
Residual		0.07	0.01	6.17	<.001	0.05	0.10
Intercept + Time	UN (1,1)	0.54	0.10	5.69	<.001	0.38	0.76
	UN (2,1)	0.02	0.01	4.31	<.001	0.01	0.03
	UN (2,2)	0.001	4.1x10^{4}	3.15	<.01	6.9x10^{4}	0.002

Estimates of Fixed Effects for Model with Reciprocal Vocal Contingency as Predictor Variable

Parameter	Estimate	*SE*	*df*	*t*	*p*	Lower CI	Upper CI
Intercept	-0.02	0.18	110.19	-0.11	.91	-0.37	0.33
Time	0.11	0.01	77.42	20.81	<.001	0.10	0.12
Reciprocal vocal contingency	0.13	0.73	84.55	0.17	.86	-1.33	1.59

Covariance Parameters for Model with Reciprocal Vocal Contingency as Predictor Variable

Parameter		Estimate	*SE*	Wald Z	*p*	Lower CI	Upper CI
Residual		0.07	0.01	6.17	<.001	0.05	0.10
Intercept + Time	UN (1,1)	0.54	0.10	5.68	<.001	0.39	0.77
	UN (2,1)	0.02	0.01	4.29	<.001	0.01	0.03
	UN (2,2)	0.001	4.1×10^{4}	3.15	<.01	6.9×10^{4}	0.002

Appendix D

Growth Curve Model Results for Incremental Validity Analyses

Research Question 4:

Estimates of Fixed Effects for Model with Only Total Vocalizations

Parameter	Estimate	*SE*	*df*	*t*	*p*	Lower CI	Upper CI
Intercept	0.21	0.09	119.83	2.36	<.01	0.03	0.38
Time	0.11	0.01	167.02	20.51	<.001	0.10	0.12
Total vocalizations	0.26	0.07	88.57	3.74	<.001	0.12	0.40

Covariance Parameters for Model with Only Total Vocalizations

Parameter	Estimate	*SE*	Wald Z	*p*	Lower CI	Upper CI
Residual	0.16	0.02	9.08	<.001	0.13	0.20
Intercept	0.26	0.05	5.36	<.001	0.18	0.38

Estimates of Fixed Effects for Model with Total Vocalizations and Communicative Use

Parameter	Estimate	*SE*	*df*	*t*	*p*	Lower CI	Upper CI
Intercept	0.69	0.08	117.88	8.26	<.001	0.53	0.86
Time	0.11	0.01	168.65	20.59	<.001	0.10	0.12
Total vocalizations	-0.01	0.06	89.15	-0.25	.80	-0.13	0.10
Communicative use composite	0.91	0.10	85.43	9.29	<.001	0.72	1.11

Covariance Parameters for Model with Total Vocalizations and Communicative Use

Parameter	Estimate	*SE*	Wald Z	*p*	Lower CI	Upper CI
Residual	0.16	0.02	9.11	<.001	0.13	0.20
Intercept	0.10	0.03	4.14	<.001	0.06	0.17

Estimates of Fixed Effects for Model with Total Vocalizations and Complexity

Parameter	Estimate	*SE*	*df*	*t*	*p*	Lower CI	Upper CI
Intercept	0.55	0.08	121.01	6.64	<.001	0.39	0.71
Time	0.11	0.01	167.74	20.60	<.001	0.10	0.12
Total vocalizations	-0.11	0.07	87.00	-1.50	.14	-0.25	0.04
Complexity composite	0.60	0.08	84.71	7.74	<.001	0.44	0.75

Covariance Parameters for Model with Total Vocalizations and Complexity

Parameter	Estimate	*SE*	Wald Z	*p*	Lower CI	Upper CI
Residual	0.16	0.02	9.09	<.001	0.13	0.20
Intercept	0.13	0.03	4.49	<.001	0.09	0.20

Estimates of Fixed Effects for Model with Total Vocalizations and ACPU-C+V

Parameter	Estimate	*SE*	*df*	*t*	*p*	Lower CI	Upper CI
Intercept	0.38	0.10	103.67	3.62	<.001	0.17	0.59
Time	0.11	0.01	163.29	20.28	<.001	0.10	0.12
Total vocalizations	0.27	0.07	87.07	3.93	<.001	0.13	0.40
ACPU-C+V	0.24	0.09	86.67	2.69	<.01	0.06	0.42

Note. ACPU-C+V = average count per utterance – consonants + vowels (Woynaroski et al., 2017; Xu, Richards, & Gilkerson, 2014).

Covariance Parameters for Model with Total Vocalizations and ACPU-C+V

Parameter	Estimate	*SE*	Wald Z	*p*	Lower CI	Upper CI
Residual	0.16	0.02	8.98	<.001	0.13	0.20
Intercept	0.24	0.05	5.25	<.001	0.17	0.35

Note. ACPU-C+V = average count per utterance – consonants + vowels (Woynaroski et al., 2017; Xu, Richards, & Gilkerson, 2014).

Estimates of Fixed Effects for Model with Total Vocalizations and Reciprocal Vocal Contingency

Parameter	Estimate	*SE*	*df*	*t*	*p*	Lower CI	Upper CI
Intercept	-0.19	0.21	93.85	-0.87	.39	-0.61	0.24
Time	0.11	0.01	162.96	20.27	<.001	0.10	0.12
Total vocalizations	0.25	0.07	87.73	3.49	.001	0.11	0.39
Reciprocal vocal contingency	1.79	0.88	86.89	2.03	<.05	0.04	3.53

Covariance Parameters for Model with Total Vocalizations and Reciprocal Vocal Contingency

Parameter	Estimate	*SE*	Wald Z	*p*	Lower CI	Upper CI
Residual	0.16	0.02	8.97	<.001	0.13	0.20
Intercept	0.25	0.05	5.25	<.001	0.17	0.36

Estimates of Fixed Effects for Model with Only Number of Child Speech-Related Vocalizations

Parameter	Estimate	*SE*	*df*	*t*	*p*	Lower CI	Upper CI
Intercept	-0.47	0.13	95.74	-3.57	.001	-0.72	-0.21
Time	0.11	0.01	163.29	20.23	<.001	0.10	0.12
Number of child speech-related vocalizations	2.9×10^4	6.8×10^5	83.89	4.25	<.001	1.5×10^4	4.3×10^4

Covariance Parameters for Model with Only Number of Child Speech-Related Vocalizations

Parameter	Estimate	*SE*	Wald Z	*p*	Lower CI	Upper CI
Residual	0.16	0.02	8.98	<.001	0.13	0.20
Intercept	0.25	0.05	5.29	<.001	0.18	0.37

Estimates of Fixed Effects for Model with Number of Child Speech-Related Vocalizations and Communicative Use

Parameter	Estimate	*SE*	*df*	*t*	*p*	Lower CI	Upper CI
Intercept	0.48	0.14	94.78	3.49	<.01	0.21	0.75
Time	0.11	0.01	164.67	20.39	<.001	0.10	0.12
Number of child speech-related vocalizations	9.8×10^{5}	5.2×10^{5}	83.81	1.89	.06	-5.0×10^{6}	2.0×10^{4}
Communicative use composite	0.83	0.09	83.88	9.39	<.001	0.66	1.01

Covariance Parameters for Model with Number of Child Speech-Related Vocalizations and Communicative Use

Parameter	Estimate	*SE*	Wald Z	*p*	Lower CI	Upper CI
Residual	0.16	0.02	8.99	<.001	0.13	0.20
Intercept	0.10	0.02	3.99	<.001	0.06	0.16

Estimates of Fixed Effects for Model with Number of Child Speech-Related Vocalizations and Complexity

Parameter	Estimate	*SE*	*df*	*t*	*p*	Lower CI	Upper CI
Intercept	0.35	0.15	94.78	2.30	.02	0.05	0.64
Time	0.11	0.01	163.51	20.39	<.001	0.10	0.12
Number of child speech-related vocalizations	9.5×10^{5}	6.9×10^{5}	83.77	1.62	.11	-2.2×10^{5}	2.1×10^{4}
Complexity composite	0.47	0.06	85.03	7.42	<.001	0.35	0.60

Covariance Parameters for Model with Number of Child Speech-Related Vocalizations and Complexity

Parameter	Estimate	*SE*	Wald Z	*p*	Lower CI	Upper CI
Residual	0.16	0.02	8.97	<.001	0.13	0.20
Intercept	0.13	0.03	4.44	<.001	0.08	0.20

Estimates of Fixed Effects for Model with Number of Child Speech-Related Vocalizations and ACPU-C+V

Parameter	Estimate	*SE*	*df*	*t*	*p*	Lower CI	Upper CI
Intercept	-0.29	0.16	92.12	-1.78	.08	-0.61	0.03
Time	0.11	0.01	163.45	20.21	<.001	0.10	0.12
Number of child speech-related vocalizations	2.6×10^{4}	6.9×10^{5}	84.41	3.69	<.001	1.2×10^{4}	3.9×10^{4}
ACPU-C+V	0.17	0.09	87.01	1.84	.07	-0.01	0.36

Note. ACPU-C+V = average count per utterance – consonants + vowels (Woynaroski et al., 2017; Xu, Richards, & Gilkerson, 2014).

Covariance Parameters for Model with Number of Child Speech-Related Vocalizations and ACPU-C+V

Parameter	Estimate	*SE*	Wald Z	*p*	Lower CI	Upper CI
Residual	0.16	0.02	8.98	<.001	0.13	0.20
Intercept	0.24	0.05	5.25	<.001	0.17	0.35

Note. ACPU-C+V = average count per utterance – consonants + vowels (Woynaroski et al., 2017; Xu, Richards, & Gilkerson, 2014).

Estimates of Fixed Effects for Model with Number of Child Speech-Related Vocalizations and Reciprocal Vocal Contingency

Parameter	Estimate	*SE*	*df*	*t*	*p*	Lower CI	Upper CI
Intercept	-0.58	0.20	91.49	-2.97	<.01	-0.97	-0.19
Time	0.11	0.01	163.13	20.21	<.001	0.10	0.12
Number of child speech-related vocalizations	2.6×10^4	7.9×10^5	83.46	3.29	<.01	1.0×10^4	4.1×10^4
Reciprocal vocal contingency	0.79	1.00	85.56	0.79	.43	-1.20	2.79

Covariance Parameters for Model with Number of Child Speech-Related Vocalizations and Reciprocal Vocal Contingency

Parameter	Estimate	*SE*	Wald Z	*p*	Lower CI	Upper CI
Residual	0.16	0.02	8.97	<.001	0.13	0.20
Intercept	0.25	0.05	5.27	<.001	0.17	0.36

Research Question 5:

Estimates of Fixed Effects for Model with Only Communicative Use Composite

Parameter	Estimate	*SE*	*df*	*t*	*p*	Lower CI	Upper CI
Intercept	0.69	0.08	117.42	8.29	<.001	0.53	0.86
Time	0.11	0.01	168.65	20.60	<.001	0.10	0.12
Communicative use composite	0.90	0.08	85.60	10.71	<.001	0.73	1.07

Covariance Parameters for Model with Only Communicative Use Composite

Parameter	Estimate	*SE*	Wald Z	*p*	Lower CI	Upper CI
Residual	0.16	0.02	9.10	<.001	0.13	0.20
Intercept	0.10	0.03	4.14	<.001	0.06	0.17

Estimates of Fixed Effects for Model with Communicative Use Composite and Reciprocal Vocal Contingency

Parameter	Estimate	*SE*	*df*	*t*	*p*	Lower CI	Upper CI
Intercept	0.52	0.17	95.00	2.98	<.01	0.17	0.86
Time	0.11	0.01	164.57	20.37	<.001	0.10	0.12
Communicative use composite	0.87	0.09	84.39	10.17	<.001	0.70	1.05
Reciprocal vocal contingency	0.72	0.64	87.15	1.12	.27	-0.56	1.99

Covariance Parameters for Model with Communicative Use Composite and Reciprocal Vocal Contingency

Parameter	Estimate	*SE*	Wald Z	*p*	Lower CI	Upper CI
Residual	0.16	0.02	8.99	<.001	0.13	0.20
Intercept	0.10	0.02	4.06	<.001	0.06	0.16

Estimates of Fixed Effects for Model with Only Complexity Composite

Parameter	Estimate	*SE*	*df*	*t*	*p*	Lower CI	Upper CI
Intercept	0.55	0.08	119.88	6.60	<.001	0.39	0.72
Time	0.11	0.01	167.41	20.60	<.001	0.10	0.12
Complexity Composite	0.52	0.06	86.03	8.94	<.001	0.40	0.63

Covariance Parameters for Model with Only Complexity Composite

Parameter	Estimate	*SE*	Wald Z	*p*	Lower CI	Upper CI
Residual	0.17	0.02	9.08	<.001	0.13	0.20
Intercept	0.14	0.03	4.52	<.001	0.09	0.21

Estimates of Fixed Effects for Model with Complexity Composite and Reciprocal Vocal Contingency

Parameter	Estimate	*SE*	*df*	*t*	*p*	Lower CI	Upper CI
Intercept	0.49	0.20	95.01	2.47	.02	0.10	0.88
Time	0.11	0.01	163.29	20.39	<.001	0.10	0.12
Complexity composite	0.51	0.06	85.95	8.20	<.001	0.39	0.63
Reciprocal vocal contingency	0.24	0.74	87.37	0.33	.74	-1.22	1.71

Covariance Parameters for Model with Complexity Composite and Reciprocal Vocal Contingency

Parameter	Estimate	*SE*	Wald Z	*p*	Lower CI	Upper CI
Residual	0.16	0.02	8.96	<.001	0.13	0.20
Intercept	0.14	0.03	4.49	<.001	0.09	0.21

Estimates of Fixed Effects for Model with Only ACPU-C+V

Parameter	Estimate	*SE*	*df*	*t*	*p*	Lower CI	Upper CI
Intercept	0.19	0.10	103.79	1.93	.06	-0.01	0.40
Time	0.11	0.01	163.23	20.20	<.001	0.10	0.12
ACPU-C+V	0.27	0.10	86.50	2.71	.01	0.07	0.46

Note. ACPU-C+V = average count per utterance – consonants + vowels (Woynaroski et al., 2017; Xu, Richards, & Gilkerson, 2014).

Covariance Parameters for Model with Only ACPU-C+V

Parameter	Estimate	*SE*	Wald Z	*p*	Lower CI	Upper CI
Residual	0.16	0.02	8.98	<.001	0.13	0.20
Intercept	0.29	0.05	5.43	<.001	0.20	0.42

Note. ACPU-C+V = average count per utterance – consonants + vowels (Woynaroski et al., 2017; Xu, Richards, & Gilkerson, 2014).

Estimates of Fixed Effects for Model with ACPU-C+V and Reciprocal Vocal Contingency

Parameter	Estimate	*SE*	*df*	*t*	*p*	Lower CI	Upper CI
Intercept	-0.25	0.24	90.86	-1.06	.29	-0.72	0.22
Time	0.11	0.01	163.09	20.17	<.001	0.10	0.12
ACPU-C+V	0.21	0.10	87.06	2.13	.04	0.01	0.41
Reciprocal vocal contingency	1.93	0.93	86.64	2.08	.04	0.08	3.77

Note. ACPU-C+V = average count per utterance – consonants + vowels (Woynaroski et al., 2017; Xu, Richards, & Gilkerson, 2014).

Covariance Parameters for Model with Complexity Composite and Reciprocal Vocal Contingency

Parameter	Estimate	*SE*	Wald Z	*p*	Lower CI	Upper CI
Residual	0.16	0.02	8.98	<.001	0.13	0.20
Intercept	0.27	0.05	5.36	<.001	0.19	0.39

Research Question 6:

Estimates of Fixed Effects for Model with Total Vocalizations and Number of Child Speech-Related Vocalizations

Parameter	Estimate	*SE*	*df*	*t*	*p*	Lower CI	Upper CI
Intercept	-0.20	0.16	96.04	-1.27	.21	-0.52	0.11
Time	0.11	0.01	163.29	20.29	<.001	0.10	0.12
Total vocalizations	0.19	0.07	88.21	2.66	<.01	0.05	0.34
Number of child speech-related vocalizations	$2.2x10^4$	$7.1x10^5$	85.11	3.07	<.01	$7.7x10^5$	$3.6x10^4$

Covariance Parameters for Model with Total Vocalizations and Number of Child Speech-Related Vocalizations

Parameter	Estimate	*SE*	Wald Z	*p*	Lower CI	Upper CI
Residual	0.16	0.02	8.98	<.001	0.13	0.20
Intercept	0.23	0.04	5.18	<.001	0.16	0.34

Estimates of Fixed Effects for Model with Complexity Composite and ACPU-C+V

Parameter	Estimate	*SE*	*df*	*t*	*p*	Lower CI	Upper CI
Intercept	0.58	0.09	109.20	6.38	<.001	0.40	0.76
Time	0.11	0.01	163.49	20.38	<.001	0.10	0.12
Complexity composite	0.50	0.06	85.33	8.26	<.001	0.38	0.62
ACPU-C+V	0.07	0.08	87.23	0.88	.38	-0.09	0.22

Note. ACPU-C+V = average count per utterance – consonants + vowels (Woynaroski et al., 2017; Xu, Richards, & Gilkerson, 2014).

Covariance Parameters for Model with Complexity Composite and ACPU-C+V

Parameter	Estimate	*SE*	Wald Z	*p*	Lower CI	Upper CI
Residual	0.16	0.02	8.97	<.001	0.13	0.20
Intercept	0.13	0.03	4.49	<.001	0.09	0.21

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. doi:10.1007/s10803-005-3302-5

Brian, J. A., Smith, I. M., Zwaigenbaum, L., & Bryson, S. E. (2017). Cross-site randomized control trial of the Social ABCs caregiver-mediated intervention for toddlers with autism spectrum disorder. *Autism Research, 10*, 1700-1711. doi:10.1002/aur.1818

Camarata, S., & Yoder, P. (2002). Language transactions during development and intervention: Theoretical implications for developmental neuroscience. *International Journal of Developmental Neuroscience, 20*, 459-465. doi: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. doi:10.1037/h0046016

Cohen, J., & Cohen, P. (1984). *Applied multiple regression*. Mahwah, NJ: Erlbaum.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52*, 281-302.

Dykstra, J. R., Sabatos-DeVito, M. G., Irvin, D. W., Boyd, B. A., Hume, K. A., & Odom, S. L. (2013). Using the Language Environment Analysis (LENA) system in preschool classrooms with children with autism spectrum disorders. *Autism, 17*, 582-594. doi:10.1177/1362361312446206

Enders, C. K. (2010). *Applied missing data analysis*. New York, NY: 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. doi: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, MD: Brookes.

Gilkerson, J., & Richards, J. A. (2008). *The LENA natural language study (LTR-02-2)*. Boulder, Colorado: LENA Foundation.

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, 100*, 8030-8035. doi: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. doi:10.1111/j.1467-9280.2008.02117.x

Goldstein, M. H., & West, M. J. (1999). Consistent responses of human mothers to prelinguistic infants: The effect of prelinguistic repertoire size. *Journal of Comparative Psychology, 113*, 52. doi:10.1037/0735-7036.113.1.52

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. doi:10.1177/105381510602800306

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. doi:10.1177/0165025406071914

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. doi:10.1901/jeab.1980.34-297

Harbison, A., Woynaroski, T., Warlaumont, A., Wade, J., Tapp, J. & Yoder, P. (2018). A new measure of vocal contingency in young children with autism spectrum disorder. *Autism Research, 11*, 903-915*.* doi:10.1002/aur.1942

Howlin, P. (2000). Outcome in adult life for more able individuals with autism or Asperger syndrome. *Autism, 4*, 63-83. doi:10.1177/1362361300004001005

Jakobson, R. (1968). Child language, aphasia, and language universals. The Hague: Mouton. [Originally published 1941]

Koopmans-van Beinum, F. J., Clement, C. J., & Van Den Dikkenberg-Pot, I. (2001). Babbling and the lack of auditory speech perception: A matter of coordination? *Developmental Science, 4*, 61-70.

Krantz, D. H. (1999). The null hypothesis testing controversy in psychology. *Journal of the American Statistical Association, 94*, 1372-1381. doi:10.1080/01621459.1999.10473888

LENA Research Foundation. (2015). 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. (2016). The relative accuracy and interpretability of five sequential analysis methods: A simulation study. *Behavior Research Methods, 48*, 1482-1491. doi:10.3758/s13428-015-0661-5

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. doi: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. doi: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.

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

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)

McCune, L., & Vihman, M. M. (2001). Early phonetic and lexical developmenta productivity approach. *Journal of Speech, Language, and Hearing Research, 44*, 670-684. doi:10.1044/1092-4388(2001/054)

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. doi:10.1016/j.ridd.2017.11.010

Miller, J. & Chapman, R. (2016) Systematic Analysis of Language Transcripts (Version 16) [Computer software]. Madison, WI: SALT Software LLC.

Mullen, E. M. (1995). *Mullen Scales of Early Learning* (AGS ed.). Los Angeles, CA: 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. doi:10.1111/infa.12090

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. doi: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, 107*, 13354-13359. doi:10.1073/pnas.1003882107

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. doi:10.1007/s10803-014-2047-4

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. doi:10.1002/aur.12

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. doi:10.1044/1092-4388(2012/11/-004)

Rankine, J. M. (2016). *Evaluating an objective measure of language in minimally verbal autism: automated Language ENvironment Analysis (LENA) in Phelan-McDermid Syndrome* (Master’s thesis). Retrieved from ProQuest Dissertations and Theses Global database. (UMI No. 10100456)

Reynell, J., & Gruber, C. (1990). *Reynell developmental language scales-U.S. edition*. Los Angeles, CA: Western Psychological Serivces.

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.

Rogers, S. J., & Dawson, G. (2009). *Early Start Denver Model for young children with autism: Promoting language, learning, & engagement*. New York, NY: Guilford Press.

Rvachew, S., Mattock, K., Polka, L., & Ménard, L. (2006). Developmental and cross-linguistic variation in the infant vowel space: The case of Canadian English and Canadian French. *The Journal of the Acoustical Society of America, 120*, 2250-2259. doi:10.1121/1.2266460

Sandbank, M. S., & Yoder, P. (2014). Measuring representative communication in 3-year-olds with intellectual disabilities. *Topics in Early Childhood Special Education, 34*, 133-141. doi:10.1177/0271121414528052

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. doi.org/10.1023/A:1005531501155.

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. doi:10.1007/s10803-010-1117-5

Shumway, S., & Wetherby, A. M. (2009). Communicative acts of children with autism spectrum disorders in the second year of life. *Journal of Speech, Language, and Hearing Research, 52*, 1139-1156. doi:10.1044/1092-4388(2009/07-0280)

Siegel, G. M., Pick, H. L. , Jr., & Garber, S. R. (1984). Auditory feedback and speech development. In H. W. Reese (Ed.), *Advances in child development and behavior* (Vol. 18, pp. 49-77). New York, NY: Academic.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York, NY: 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.

Stoel-Gammon, C. (1989). From babbling to speech: Some new evidence. Paper presented at the Child Phonology Conference, Evanston, IL.

Stoel-Gammon, C. (1991). Normal and disordered phonology in two-year-olds. *Topics in Language Disorders, 11*, 21-32. doi:10.1097/00011363-199111040-00005

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, MA: 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.

VanDam, M., & Silbert, N. H. (2016). Fidelity of Automatic Speech Processing for Adult and Child Talker Classifications. *PloS ONE, 11*, e0160588. doi:101371/journal.pone.0160588

Vihman, M. M. (2017). Learning words and learning sounds: Advances in language development. *British Journal of Psychology, 108*, 1-27. doi:10.1111/bjop.12207

Vihman, M. M., Macken, M. A., Miller, R., Simmons, H., & Miller, J. (1985). From babbling to speech: A re-assessment of the continuity issue. *Language, 61*, 397-445. doi:10.2307/414151

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. doi:10.1177/0956797614531023

Watt, N., Wetherby, A., & Shumway, S. (2006). Prelinguistic predictors of language outcome at 3 years of age. *Journal of Speech, Language, and Hearing Research, 49*, 1224-1237. doi:10.1044/1092-4388(2006/088)

West, M. J., & Rheingold, H. L. (1978). Infant stimulation of maternal instruction. *Infant Behavior and Development, 1*, 205-215. doi:10.1016/S0163-6383(78)80031-9

Wetherby, A. M., Cain, D. H., Yonclas, D. G., & Walker, V. G. (1988). Analysis of Intentional communication of normal children from the prelinguistic to the multiword stage. *Journal of Speech, Language, and Hearing Research, 31*, 240-252. doi:10.1044/jshr.3102.240

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

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. doi: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. doi:10.1044/jshd.5402.148

Wilkinson, L. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist, 54*, 594-605. doi:10.1037/0003-066X.54.8.594

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. 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.

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. doi: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. doi:10.1044/2014_JSLHR-S-13-0037

Xu D., Yapanel U., & Gray S. (2009). Reliability of the LENA™ language environment analysis system in young children's natural home environment. Retrieved from http://www.lenafoundation.org/TechReport.aspx/Reliability/LTR-05-2

Xu, D., Yapanel, U., Gray, S., & Baer, C. T. (2008). The LENA™ language environment analysis system: The interpreted time segments (ITS) file. Retrieved from http://www.lenafoundation.org/TechReport.aspx/ITS_File/LTR-04-2.

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, MD: Paul H. Brooks Publishing Company.

Yoder, P. J., Oller, D. K., Richards, J. A., Gray, S., & Gilkerson, J. (2013). Stability and validity of an automated measure of vocal development from day-long samples in children with and without autism spectrum disorder. *Autism Research, 6*, 103-107. doi:10.1002/aur.127

Yoder, P., Watson, L., & 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. doi:10.1007/s10803-014-2286-4

Zimmerman, I.L., Steiner, V.G., & Pond, R.E. (2002) *The Preschool Language Scale – Fourth Edition.* San Antonio, TX: The Psychological Corporation.