ACOUSTIC PROPERTIES OF EARLY VOCALIZATIONS IN INFANTS WITH FRAGILE X SYNDROME

by

Lisa M. Rague

A Thesis

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Master of Science

Department of Psychological Sciences

West Lafayette, Indiana

December 2018

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. Bridgette L. Tonnsen, Chair

Department of Psychological Sciences

Dr. Donald R. Lynam

Department of Psychological Sciences

Dr. Amanda H. Seidl

Department of Speech, Language and Hearing Sciences

Approved by:

Dr. David Rollock

Head of the Graduate Program

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Author: Rague, Lisa M. MS
Institution: Purdue University
Degree Received: December 2018
Title: Acoustic Properties of Early Vocalizations in Infants With Fragile X Syndrome
Committee Chair: Bridgette L. Tonnsen

Fragile X syndrome (FXS) is a neurogenetic syndrome characterized by cognitive impairments and high rates of autism spectrum disorder (ASD). FXS is often used as a model for exploring mechanisms and pathways of symptom expression in ASD due to the high prevalence of ASD in this population and the known single-gene cause for ASD in FXS. Early vocalization features – including volubility, canonical complexity, vocalization duration and vocalization pitch – have shown promise in detecting ASD in idiopathic ASD populations but have yet to be extensively studied in a population with a known cause for ASD, such as FXS. The present study characterizes early vocalization features in FXS, demonstrating how these features are associated with language ability and ASD outcomes, as well as highlighting how these features in FXS may diverge from patterns observed in typically developing (TD) populations. We coded vocalization features during a standardized child-examiner interaction in 39 nine-month-old infants (22 FXS, 17 TD) who were then followed up at 24 months to determine developmental and clinical outcomes. Although many findings did not reach statistical significance in this small sample, our results provide preliminary evidence that infants with FXS may demonstrate patterns of associations with 24-month language outcomes that diverge from those observed in typical development, and that certain vocalization features may be associated with later ASD outcomes in the FXS group. These findings warrant more research exploring these features as potential early markers of ASD in FXS. Characterizing the associations of early vocalization features with ASD outcomes in FXS can inform mechanisms of ASD development that can then be tested broadly with other etiologically-distinct populations at risk for ASD. Thus, further characterization of these early vocalization features in typical and atypical development may lead to improved early identification methods, treatment approaches, and overall well-being of individuals in the ASD population.

INTRODUCTION

Fragile X syndrome (FXS) is a neurogenetic disorder that is the leading known single-gene cause of autism spectrum disorder (ASD). Because ASD is so prevalent in FXS, and it has a well-known single-gene cause, FXS is often used as a model for exploring mechanisms and pathways of symptom expression in ASD (McCary & Roberts, 2013). Symptoms in the domain of social communication are of particular interest, as these features are core areas of deficit in ASD (APA, 2013), and individuals with FXS often show impairment in this domain, as well (Finestack, Richmond, & Abbeduto, 2009). As such, identifying early markers that predict these impairments in later development is an important endeavor that could facilitate routing high-risk children to appropriate and targeted treatments. Current efforts are focused on exploring these early markers in ASD, but little research has investigated these features in FXS to potentially inform mechanisms that predict later social communication outcomes. We address this gap by examining a model of early social communication in FXS, with two main goals: 1) to compare patterns of early vocalization features, including volubility, complexity, average vocalization duration, duration range, average pitch, and pitch range, in FXS with those observed in typically developing populations and 2) to examine how these features are related to outcomes later in development, including language abilities and ASD symptomatology. By characterizing early vocalization features in FXS, we aim to identify patterns and early markers that can be used to inform early identification and treatment efforts, as well as clarify potential etiological pathways for disorders, such as ASD, that are common in this and other genetic syndrome populations.

Early Social Communication in Fragile X Syndrome

FXS is a genetic disorder caused by an atypical expansion of CGG repeats on the X chromosome. This expansion activates a pathway that ultimately leads to reduced production of fragile X mental retardation protein (FMRP; Bhakar, Dölen, & Bear, 2012). This protein has been linked to cognitive functioning and synaptic plasticity, and its reduced production leads to cognitive deficits and often intellectual disability, particularly in males who do not have a second X chromosome to buffer negative effects (Santoro, Bray, & Warren, 2012). Autism spectrum disorder is a highly comorbid disorder in FXS, with about 30-50% of individuals with FXS also receiving a diagnosis of ASD (Hall, Lightbody, & Reiss, 2008; Harris et al., 2008). Furthermore, many individuals with FXS who do not meet the full criteria for ASD nonetheless exhibit ASD-like symptoms (Abbeduto, McDuffie, & Thurman, 2014; Hall et al., 2008), highlighting the broader association of ASD features with the FXS phenotype.

Given the high rates of comorbidity and the presence of subthreshold ASD symptoms in FXS, clarifying the complex relationship of ASD with FXS has been the focus of recent research efforts. There has been some debate as to whether ASD coupled with FXS represents a truly comorbid disorder or is simply a part of the FXS phenotype more broadly. Growing evidence has emerged from this debate suggesting that the ASD symptom profile in FXS is distinct from the symptom profile in idiopathic ASD (i.e., ASD without a known genetic marker; Budimirovic & Kaufmann, 2011). Fragile X syndrome with comorbid ASD is associated with lower levels of ASD symptoms pertaining to social interaction and reciprocity, and higher levels of symptoms of restricted and repetitive interests compared to the symptom profile of idiopathic ASD

(Hall, Lightbody, Hirt, Rezvani, & Reiss, 2010; McDuffie, Thurman, Hagerman, & Abbeduto, 2015). Within FXS, there is evidence that social behaviors can function to differentiate between individuals with FXS with ASD and those with FXS without ASD. For example, a number of studies have identified low adaptive socialization skills as a significant predictor of ASD in FXS (Budimirovic et al., 2006; Kau et al., 2004; Kaufmann et al., 2004). Furthermore, FXS with comorbid ASD is uniquely associated with lower rates of behaviors that indicate social approach or "warming up" to new people such as use of eye contact, warm facial expressions, and physical approach, compared to those with FXS and no comorbid ASD diagnosis (Roberts, Weisenfeld, Hatton, Heath, & Kaufmann, 2007). Thus, social features may be key in detecting comorbid ASD diagnoses within FXS.

While there are some social behaviors that differentiate individuals with FXS with and without comorbid ASD diagnoses, other aspects of social communication are often impaired in FXS regardless of ASD status. Some deficits in the pre-linguistic components of language in FXS become apparent early in development. For example, young children with FXS have relative weaknesses in early gesture use (Rague, Caravella, Tonnsen, Klusek, & Roberts, 2018; Roberts, Mirrett, Anderson, Burchinal, & Neebe, 2002) and may use gestures and other forms of early social communication for a limited range of functions (Marschik et al., 2014). FXS is also associated with delayed expressive and receptive language skills (Finestack et al., 2009). Indeed, individuals with FXS often remain in the pre-linguistic phase of language development (i.e., communicating nonverbally or with limited use of meaningful words as opposed to using language functionally) for longer than typically developing individuals, in some cases even into the

teenage and adult years (Brady, Skinner, Roberts, & Hennon, 2006; Levy, Gottesman, Borochowitz, Frydman, & Sagi, 2006). For those who have mastered the functional use of language, pragmatic language can be particularly compromised, even when controlling for overall cognitive ability (Klusek, Martin, & Losh, 2014; Losh, Martin, Klusek, Hogan-Brown, & Sideris, 2012). Thus, social communication is an area of particular concern for individuals with FXS across multiple domains throughout the lifespan.

Features of Early Vocalizations Associated With ASD Risk

Understanding the mechanisms by which social communication impairments occur is important for early identification as well as the development of treatments that target core processes leading to impairment. Because FXS has a known genetic cause, characterizing social communication impairments and their predictors in this population can inform our understanding of the underlying mechanisms and markers of these impairments in FXS, which could possibly extend to other high-risk groups for which no genetic cause is known. Recently, early vocalization features have been identified as potential early markers for ASD. While little research has focused on features of early vocalization in FXS, there is growing literature supporting the utility of these features in identifying individuals who go on to receive a diagnosis of ASD in non-FXS samples. In particular, early vocalization features such as volubility, complexity, average vocalization duration, duration range, average pitch, and pitch range have been previously studied in typically developing populations and populations at risk for ASD. Examining these features in FXS is an important next step for understanding the unique intersection of ASD risk with potential genetic underpinnings into which other populations studied thus far have provided limited insight.

Volubility

Volubility refers to the overall amount or rate of vocalization. As typically developing (TD) infants age, they begin to vocalize more frequently, producing more speech-like syllables (Oller, Eilers, Steffens, Lynch, & Urbano, 1994). Furthermore, by 6 months of age, infants tend to recognize the social quality of their vocalizations and have been shown to increase volubility in order to re-engage an adult who has stopped interacting with them (Franklin et al., 2013; Goldstein, Schwade, & Bornstein, 2009). Findings on the impact of certain risk factors on an infant's level of volubility are somewhat mixed. On one hand, infants living in low socioeconomic status households have consistently been shown to have lower vocalization rates compared to middle or high socioeconomic status peers (Oller et al., 1994). On the other hand, some risk factors expected to affect developmental trajectories, such as preterm birth, do not appear to affect levels of volubility relative to those of TD infants (Oller et al., 1994). Therefore, further research is needed to clarify the impact of other developmental risk factors, such as having a neurogenetic syndrome like FXS, on levels of volubility.

Though previous findings have determined that overall language development is delayed in FXS (for a review, see Abbeduto, Brady, & Kover, 2007) and that these delays are linked to level of ASD symptomatology in older children (e.g., Martin, Losh, Estigarribia, Sideris, & Roberts, 2013), little research has focused on rate of vocalization in individuals with FXS, particularly in infants. Indeed, only one small study to date has examined volubility in infants with FXS. In this study, 9- to 12-month-olds with FXS (n=10) were shown to have lower overall volubility than age-matched TD peers (n=14; Belardi et al., 2017). However, the relationship of volubility with ASD status in FXS

remains unclear because the FXS group in this study only included infants who did not go on to receive a comorbid ASD diagnosis.

In contrast, volubility in idiopathic ASD has been more extensively studied. As early as 9 months of age, children later diagnosed with ASD use speech-related utterances at lower rates than TD children (Patten et al., 2014; Plumb & Wetherby, 2013; Schoen, Paul, & Chawarska, 2012; Warren et al., 2010). The rates of overall speech vocalizations observed in 18- to 36-months-olds with ASD are consistent with those of age-matched minimally verbal children with developmental delay (Sheinkopf, Iverson, Rinaldi, & Lester, 2012), as well as those observed in younger TD infants matched on language abilities (i.e., 11- to 13-month-olds; Schoen et al., 2012). These findings suggest that reduced volubility may be related to broader developmental delays typical in ASD. Given the established state of volubility patterns in ASD, this information can be used to inform hypotheses about the implications of levels of volubility on ASD risk in FXS. Specifically, volubility may be less likely to predict later ASD features and may instead be closely related to developmental functioning.

Vocalization Complexity

In typical development, speech is acquired in a specific developmental sequence. During the first year the sequence begins with reflexive vocalizations, such as crying, from birth. At around 4 to 6 months of age (Oller & Eilers, 1988), infants begin to use vocalizations that sound slightly more speech-like, and that typically consist of various vowel sounds (e.g. “ah,” “eee-oh,”) or elongated consonants (e.g. “mm”). These types of vocalizations are often referred to as “pre-canonical” vocalizations (Oller, Eilers, & Basinger, 2001). Most typically developing infants then begin to produce vocalizations

that contain clearer consonants and vowels in “canonical” syllables around 7-10 months of age (Morgan & Wren, 2018). At around 12 months of age, most children begin to use meaningful words or word approximations (Vihman & Vihman, 2011), though there is considerable variability in the emergence of this skill across individuals. Delays in the onset of canonical babbling have been shown to predict speech delays (Oller, Eilers, Neal, & Schwartz, 1999), and this developmental sequence is often atypical or delayed in high-risk populations. For instance, on average, the onset of canonical babbling is delayed in infants with Down syndrome (DS) relative to TD infants (Lynch, Oller, Steffens, & Levine, 1995), and ratios of canonical babbling usage increase at slower rates in infants later diagnosed with language delay (Xu, Richards, & Gilkerson, 2014).

Belardi et al. (2017) is the only study to date that has examined the development of early vocalization complexity in infants with FXS (n=10). Results revealed that 9- to 12-month-olds with FXS used lower ratios of canonical syllables than age-matched TD infants. Again, as this study did not explore the implications of ASD on these findings in FXS, further research is needed to determine the clinical implications and predictive utility of vocalization complexity in FXS. Literature on vocalization complexity in ASD is slightly more established, and can provide a starting point for research on this early vocalization feature in FXS. The average child with ASD has developed less complex speech (i.e., is not yet using canonical syllables) than age-matched TD children (Patten et al., 2014; Schoen et al., 2012) and tends to use lower ratios of canonical syllables than both TD infants and infants later diagnosed with language delays (Patten et al., 2014; Xu et al., 2014). Overall, both infants later diagnosed with ASD and those with FXS tend to use lower ratios of canonical syllables, though additional research is needed to

corroborate this finding in FXS and to examine the impact of ASD on ratios of canonical syllable usage.

Average Vocalization Duration

The average duration of a speech-like syllable in infants ranges from 110 to 600 ms, while non-speech syllables are slightly longer (Oller et al., 2010). Thus, average duration of the syllable often decreases with age, as infants begin to develop more complex speech and use higher rates of speech-like syllables. Oller et al. (2010) explored the ratio of use of short (110-250 ms), medium (250-600 ms), long (600-900 ms), and extra-long (900-3000 ms) vocalizations, and found that rate of short, medium and long vocalizations increased with age, while the rate of extra-long vocalizations decreased with age. A similar pattern emerged in this study for infants with language delay, where ratio of short and medium vocalization usage increased with age and use of long and extra-long vocalizations ratios decreased with age (Oller et al., 2010). However, in a study of 2- to 12-month-olds, infants with DS produced vocalizations with longer syllables than age-matched TD infants (Lynch, Oller, Steffens, & Buder, 1995). This suggests that infants with known cognitive impairments that place them at risk for language delays, such as infants with DS, may use longer syllables than their TD peers.

There have been no studies to date examining the duration of syllables in infants with FXS relative to TD infants. However, duration of vocalizations has been studied as an element that differentiates vocalizations of children with ASD from those of TD children, with mixed findings across studies. In a recent meta-analysis of acoustic patterns of speech in individuals with ASD, 7 studies reported longer duration, 1 reported shorter duration, and 6 reported null findings (Fusaroli, Lambrechts, Bang, Bowler, &

Gaigg, 2017). A similar pattern of mixed results is seen in studies of vocalization duration in very young populations. One study reports longer vocalizations in children with ASD aged 16 to 48 months compared to those of TD children (Warren et al., 2010), while other studies report nonsignificant differences (Brisson, Martel, Serres, Sirois, & Adrien, 2014; Quigley, McNally, & Lawson, 2016). These findings suggest that vocalizations of children with ASD tend to be longer than those of their TD peers, but further research is needed to determine whether this difference is meaningful and relevant to the early identification of ASD.

Pitch

Pitch is an important aspect of prosody, which is defined as the dynamic variations in the suprasegmental qualities of speech that often convey pragmatic information (Wilson & Wharton, 2006). Research has shown that very young infants begin to systematically modulate pitch during the pre-linguistic phase (Snow & Balog, 2002). Indeed, there is evidence that TD infants as young as 3 months of age may use different pitch qualities for vocalizations with different pragmatic functions (Gratier & Devouche, 2011; Marcos, 1987). For example, infants used different pitch contours when imitating their mothers' vocalizations as opposed to the pitch contours they used when repeating their own preceding vocalizations, presumably reflecting the different pragmatic intentions behind each type of vocalization (Gratier & Devouche, 2011). Thus, pitch quality can impact social interaction, even at a very young age.

Two pitch features have previously been studied with regard to infant vocalizations: average pitch and pitch range. Average pitch refers to the average pitch with which an infant vocalizes, while pitch range refers to the difference between

absolute minimum and absolute maximum pitch the infant uses across vocalizations. These pitch features have yet to be explored in infants with FXS, though one study does suggest that ultrasonic vocalizations of the Fmr1 knockout mouse model were higher pitched and had wider pitch ranges than the wild-type mouse (Roy, Watkins, & Heck, 2012). A number of recent studies have explored whether individuals with ASD demonstrate atypicality in the prosodic properties of their vocalizations as infants. In the previously mentioned meta-analysis, the average pitch and pitch range of speech-related vocalizations were found to be significantly different in ASD versus TD populations (Fusaroli et al., 2017), with vocalizations of children with ASD generally tending to have higher pitch and a wider pitch range (Filipe, Frota, Castro, & Vicente, 2014; Sharda et al., 2010). Notably, only two studies of pitch features of speech-related vocalizations in younger populations at risk for ASD were identified. The first study retrospectively analyzed home videos of infants 6 months of age and younger who were later diagnosed with ASD (Brisson et al., 2014), while the second conducted prospective analyses of infant vocalizations produced during a naturalistic mother-infant interaction at 12 months of age (Quigley et al., 2016). Both studies concluded that average pitch and pitch range did not differ between infants later diagnosed with ASD and those with typical development. Importantly, reliable markers that predict the later diagnosis of ASD have not yet been identified before 12 months of age for a variety of methodological and theoretical reasons (Zwaigenbaum et al., 2015). Thus, further research is needed to clarify the potential of pitch features of early vocalizations as an early marker of ASD in very young infants.

Summary and the Present Study

Features of early vocalizations have shown promise as potential early markers for ASD but have been largely understudied in FXS, a population uniquely positioned to inform underlying genetic mechanisms of social communication impairment and ASD. Evidence from a single study suggests that vocalization volubility and complexity are atypical in FXS; however, no research to date has explored additional features of early vocalizations in FXS, such as average vocalization duration, duration range, average pitch, or pitch range. In contrast, literature on features of early vocalizations in ASD is more extensive. Volubility, complexity, duration and pitch qualities have all shown promise in distinguishing those who have ASD from those with typical development. Together, these studies suggest that early vocalizations in ASD are atypical, but little is known regarding how these features manifest in FXS and how ASD may impact early vocalizations in this population. However, findings in the ASD literature can be leveraged in the exploration of early vocalization patterns in infants with FXS and associations between infant vocalizations and later ASD and language outcomes. To this end, the present study will examine the association of acoustic properties of early vocalizations with autism and language outcomes in infants with FXS.

METHOD

Participants

Raw videos and psychological assessment data were drawn from a previously published study of early markers of ASD in high-risk populations (R01MH090194; PI: Jane Roberts). In this study, children seen in-person at 9, 12, and 24 months of age completed a range of measures assessing temperament, ASD risk, and developmental skills, including a child-examiner interaction task at the 9- and 12-month assessments. In the present study, we focus on interactions completed at the 9-month assessment to capture the earliest time point available at which infants may be expected to demonstrate emerging use of canonical syllables (Oller, Eilers, Neal, & Schwartz, 2006). Of the original 43 infants with FXS assessed in the original study, a subsample of 22 infants with FXS were included who 1) have complete data from the 9-month assessment, including developmental testing and the child-examiner interaction and 2) have completed an ADOS-T and were provided a clinical diagnosis of ASD (FXS-ASD, n = 10; 6 females), or Non-ASD (FXS-O; n = 12, 4 females) at their final assessment. To inform patterns of vocalization features that converge or diverge with those observed in typical development, we also included a control group of 17 TD infants (5 female), all of whom were determined to show no clinical features of ASD or developmental delays based on clinical evaluations at their final assessment. The FXS group was slightly older than the TD group at the 9- and 24-month assessments, and the FXS-ASD group was slightly older than the FXS-O group at the 24-month assessment (Table 1). Demographic information is presented in Table 2.

Methodological details for the original project, including inclusion and exclusion criteria, were previously published in Roberts, Tonnsen, McCary, Caravella, and Shinkareva (2016). In brief, infants with neurological conditions or born prematurely (<37 weeks) were excluded from the study. FXS status was confirmed by genetic report. Procedures for secondary analyses were approved by the Institutional Review Board at Purdue University, and initial study and consent procedures were approved by the University of South Carolina Institutional Review Board.

Measures

Standardized Child-Examiner Interaction

Vocalization features were coded during the Autism Observation Scale for Infants (AOSI; Bryson, Zwaigenbaum, McDermott, Rombough, & Brian, 2008), a semi-structured, standardized child-examiner interaction used to evaluate autism risk in infants 6 to 18 months of age. During the AOSI, a trained examiner interacts with the infant for 15 to 20 minutes using pre-designated activities and presses to observe the infant's social behavior and any signs of known early risk markers for ASD. The AOSI begins with the examiner testing the infant's ability to attend to a stimulus across their visual field, and then the ability to disengage their visual attention from a stimulus when a novel stimulus is presented. Next, the infant participates in a "free play" interaction, where a series of age-appropriate toys (e.g., cloth book, rattles, rubber blocks) are presented to the infant while the examiner attempts to engage the infant in social play with the toys. After 3 to 5 minutes, the examiner removes the toys and initiates a social routine with the infant (e.g., peekaboo, tickling game). Next, the examiner demonstrates three simple actions and observes the infant's ability to imitate these actions. The AOSI then concludes with

another 3- to 5-minute “free play” interaction. Throughout these activities, the infant is given ample opportunity to both produce spontaneous vocalizations, as well as to respond to the examiner’s vocalizations. In this study, all AOSIs were administered in the infant’s home.

Nonverbal Mental Age

Nonverbal Mental Age (NVMA) was calculated for each participant based on scores from the Mullen Scales of Early Learning (MSEL; Mullen, 1995). The MSEL is a measure of developmental ability used with children ages 0 through 68 months, collected at the 9- and 24-month assessments in the original study. Five domains of development are measured: Gross Motor, Visual Reception, Fine Motor, Receptive Language, and Expressive Language. In this study, 9-month and 24-month NVMA were calculated by averaging the Visual Reception and Fine Motor age equivalents (Munson et al., 2008).

Language Outcomes

Receptive and expressive language outcomes were measured using age equivalents from the Expressive Language and Receptive Language scales on the MSEL (MSEL-EL AE and MSEL-RL AE, respectively) collected at the 24-month assessment. Age equivalents were used instead of standard scores to avoid the floor effects and limited variability often observed when using standard scores in severely developmentally delayed populations like FXS.

Autism Symptoms and Diagnoses

Autism symptom severity was measured using the Autism Diagnostic Observation Schedule – 2nd Edition, Toddler Module (ADOS-T; Lord et al., 2012), a semi-structured standardized interaction used to observe behaviors that can be indicative

of ASD in children ages 18-30 months who are not yet using two-word phrases. During the ADOS-T, a trained examiner engages the child in a series of planned social situations and rates the child's behavior across various domains related to characteristics of ASD. A subset of these ratings is used to calculate an overall total score, which can then be subsetted and converted to Calibrated Severity Scores that can be compared across ages. The Social Affect Calibrated Severity Score (ADOS-T SA CSS; Esler et al., 2015) from the ADOS-T collected at the 24-month assessment was used as a continuous measure of atypical social behavior potentially related to ASD. Higher ADOS-T SA CSS scores indicate higher levels of ASD symptomatology. We chose the Social Affect severity score based on evidence that social communication symptoms tend to differentiate best between individuals with and without ASD in syndromic populations, while restricted and repetitive behaviors are typically elevated in these populations regardless of ASD diagnosis (Budimirovic et al., 2006).

We also used available data on clinical classification of ASD and non-ASD status, which was determined in the original study using Clinical Best Estimate (CBE; Lord, 2012) procedures. Various data that can inform an ASD diagnosis were accumulated for each child, including the ADOS-T, the Autism Diagnostic Interview – Revised (ADI-R; Le Couteur, Lord, & Rutter, 2003) and the MSEL. These data were reviewed by a licensed psychologist and two other ADOS-reliable researchers, including at least one researcher who conducted the assessment with the child. The CBE meeting for each study participant involved watching 15 minutes of the child's ADOS-T and reviewing scores from the child's ADOS-T, ADI-R, and MSEL. Based on this information, the child was determined to have ASD, Subthreshold ASD, Non-ASD Developmental Delay, or No

Clinical Features, with the CBE team indicating the certainty with which they endorse the diagnosis for each case on a 5-point scale (0-20% certain, 20-40% certain, 40-60% certain, 60-80% certain, 80-100% certain). In this study, the CBE was used to establish the FXS-ASD and the FXS-O groups for exploratory categorical analyses.

Data Preparation

Data preparation and coding procedures are presented in Figure 1. Child-examiner interaction videos were trimmed using Pavtube Video Converter software (Version 4.8.4.0, 2016). After being clipped, videos were converted to Waveform audio files (.wav) using Pavtube. Clipped videos were then imported into ELAN (Wittenburg, Brugman, Russel, Klassmann, & Sloetjes, 2006) to create a coding file that could be read into the coding software along with the Waveform audio file.

Acoustic Coding

Coding Procedures

Coding was conducted using Praat (Boersma & Weenink, 2016). For each file, two trained coders simultaneously listened to an audio recording of the child-examiner interaction, with both coders wearing headphones to minimize background noise in the coding environment. Coding was based on audio recordings alone – videos were not referenced during coding. If a sound could not be determined to be an infant vocalization with certainty (e.g., it may have instead been background noise or a vocalization made by another person in the video), it was not coded. Similarly, if a sound was too quiet to show up in the Praat spectogram window or Praat was unable to create a pitch track of the sound, the sound was not coded. Once a sound was identified as a codable vocalization, coders marked boundaries at the beginning and end of each vocalization, using both the

waveform and spectogram windows in Praat to inform coding decisions (Figure 2; see Appendix C for detailed coding schemes). When determining whether a vocalization was composed of multiple vocalizations or one long vocalization, coders made the decision of whether to break into separate vocalizations based on whether they heard one or more syllables. Due to the interactive nature of the AOSI, it was common for non-infant sounds (e.g., examiner talking, noise from toys) to overlap with infant vocalizations. In these cases, coders placed boundaries as close to the beginning or end of the child vocalization, ignoring other sounds.

Coding Scheme

Each vocalization was assigned two codes. The first code, "Codability," was used to determine whether extraneous sounds occurred within the boundaries of the coded infant vocalization. Vocalizations with overlapping sounds, such as noise from toys or other people in the room talking, were given a code of "0," while vocalizations with no background noise were given a code of "1". Vocalizations with a Codability code of "0" were excluded from analyses involving pitch, which may be affected by extraneous noises. The second code, "Speech Type," was used to indicate the complexity of the vocalization. Three levels of speech type were coded, based on definitions established in Oller and Eilers (1988). Any vocalization that the coders identified as a cry, laugh, grunt, or squeal was coded as Non-Speech. Vocalizations that contained only vowel sounds (e.g., "eh," "ooo,"), or elongated consonant sounds (e.g., "mmm") were coded as Pre-Canonical. Finally, vocalizations that contained both a consonant and a vowel with rapid transitions between the two (e.g., "ba," "ma") were coded as Canonical.

Training and Reliability

Undergraduate coders were trained on the coding scheme by the first author. Coders first reviewed relevant literature about measuring acoustic properties of infant speech (Patten et al., 2014; Paul, Fuerst, Ramsay, Chawarska, & Klin, 2011), as well as a detailed coding manual describing data preparation procedures, coding procedures and descriptions of the coding scheme specific to this study (see Appendix C). Coders then met with the first author to review coding techniques and definitions and apply them to a practice file. After this training, coders began paired coding of files of excluded participants in order to establish reliable coding with the other coding pairs.

Reliability was calculated for the two codes (Codability and Speech Type) as well as three additional variables: Code Status (whether both coding pairs coded the vocalization), Boundary Placement (whether both coding pairs placed vocalization boundaries within 0.075 seconds of each other, averaged across the beginning and end boundaries), and Breaks (whether both coding pairs broke vocalization groups into the same number of syllables). Coding pairs coded pre-determined reliability files which were compared to other coding pairs who had coded the same reliability file, with reliability defined as 70% agreement for each of the 5 reliability variables.

Data Processing

For each participant, data for each vocalization was extracted using Praat scripts, which extracted the start and end time of each vocalization and the Speech Type and Codability of each vocalization. After these data were extracted, vocalizations occurring within 1 second of each other were identified using Excel formulas, and were combined into a single vocalization group (Paul et al., 2011). In these instances, the highest Speech

Type rating for the vocalization group was assigned to the overall vocalization. For example, if a child said “ah ah ba,” with less than 1 second between each syllable, this vocalization would be coded as a single vocalization and would be assigned a Speech Type code of 2 to give the child credit for the highest level of speech produced. Similarly, if any syllable in the group was assigned a Codability code of 0, the overall vocalization was also assigned a Codability code of 0. Thus, if the first “ah” in the vocalization from the previous example were overlapped by a toy noise, the entire vocalization would be assigned a Codability code of 0 to reflect the fact that part of the total vocalization contains overlapping noise.

After data processing, six vocalization variables were calculated for each participant: volubility, complexity, average vocalization duration, duration range, average pitch, and pitch range. Volubility was defined as the number of speech syllables per minute, calculated by dividing the total number of speech syllables by the length of the audio clip in seconds and multiplying this by 60 (Oller et al., 1994). This rate calculation accounts for the variability in the length of the child-examiner interaction, so that infants with shorter interactions did not receive lower counts due to a shorter opportunity to vocalize. Complexity was calculated as the ratio of canonical syllables over total speech syllables (Belardi et al., 2017; Patten et al., 2014). After conducting preliminary analyses, it became clear that the usage of any canonical syllables by infants in both the FXS and TD groups was rare (Table 3), leading to significant floor effects for this variable. Thus, we also conducted analyses defining complexity as either the presence or absence of canonical syllable usage.

The average duration of vocalization was calculated for each participant by averaging the length in seconds of all speech vocalizations. The range of vocalization duration was also computed by calculating the difference between the shortest and the longest vocalization duration across all speech vocalizations per participant. Average pitch was calculated for each participant by averaging the pitch of all speech vocalizations. Pitch range was determined by calculating the difference between the minimum and maximum pitch across all speech vocalizations per participant. Only codable vocalizations were used in calculating pitch variables.

Power Analysis

Sensitivity power analyses were computed in G*Power 3.1. Due to our small sample size, we interpreted p-values between .05 and .10 as trends, and thus conducted power analyses of our ability to detect effects at the .10 alpha level. For FXS vs. TD group comparisons, based on an FXS sample size of 22 and a TD sample size of 17, we observed 80% power to detect large effect sizes (d = .71) for Wilcoxon Rank-Sum tests of mean differences. For associations with 24-month outcomes, we had 80% power to detect large effect sizes (FXS: f^2 = 0.30 for n = 22; TD: f^2 = 0.40 for n = 17) for the Spearman rank-order partial correlation analyses of the outcome measures (i.e., MSEL-EL AE, MSEL-RL AE, ADOS-T SA CSS) with each vocalization feature, controlling for 24-month level of developmental delay. Despite limited power to detect small-to-medium effects, our study addresses a much-needed area of research in the field given the low availability of data on early development in FXS. Furthermore, it will advance the field by reporting novel information about developmental and clinical outcomes in FXS that can be used to inform hypotheses and justify resources for subsequent, higher-powered

studies. However to temper potential risk of Type I error, we used best practices for presenting small syndromic data, including using non-parametric statistical tests, reporting effect sizes for all results, and including plots of raw data. We also present a summary of results that did not were not statistically significant but that did demonstrate medium-sized effects or larger as exploratory results that are candidates for further investigation in studies with larger sample sizes.

RESULTS

Analytic Plan

We first examined differences in mean levels of 9-month vocalization features between the FXS and TD groups using Wilcoxon Rank-Sum tests. We expected that the TD group would demonstrate more developmentally appropriate patterns of vocalization features, including having higher volubility and canonical syllable usage. We also expected the TD group would demonstrate shorter average duration and pitch, and less variable duration and pitch ranges based on literature examining speech in older children; however, it is unclear whether these patterns will extend to canonical or pre-canonical syllables in infants.

Next, we analyzed associations of vocalization features at 9 months with language outcomes at 24 months in the TD and FXS groups separately using Spearman rank-order semi-partial correlations of 9-month vocalization features with 24-month receptive and expressive language age equivalents, controlling for 24-month level of developmental delay. We calculated a measure of level of developmental delay by subtracting each child's 24-month NVMA from their 24-month chronological age, with higher values indicating more discrepancy between developmental skills expected for the child's age and the observed developmental skills for that child. This allows the conclusion that any observed associations of vocalization features with language outcomes are not better explained by broad developmental delays. We expected that for both groups, higher 24-month language abilities would be related to higher volubility and canonical complexity, shorter average duration and pitch, and less variable duration and pitch ranges.

Finally, we examined the association of vocalization features with autism outcomes in the FXS group by testing (1) the association of continuous autism symptom severity scores with vocalization features, and (2) differences in vocalization features based on categorical ASD diagnosis, which incorporates multiple measures of ASD and clinician judgement. We first examined continuous associations by conducting Spearman semi-partial correlations of 9-month vocalization features with ASD symptom severity at 24 months, controlling for level of 24-month developmental delay. Next, we examined differences in mean levels of vocalization features between the FXS-O and FXS-ASD groups, defined by categorical clinical best estimate rather than continuous symptom scores, to explore the effect of ASD diagnosis on early vocalization features in the FXS group. The TD group was excluded due to expected low variability in autism symptoms. We predicted that the FXS-ASD group and those with higher ADOS-T SA CSS would demonstrate lower volubility and canonical complexity, longer average duration and higher average pitch, and more variable duration and pitch ranges.

We conclude with a discussion of associations that did not reach statistical significance but nevertheless demonstrated medium-sized effects or larger. These exploratory results may suggest areas for further exploration in studies with larger samples.

Comparison of FXS and TD Vocalization Features

To determine whether vocalization features in FXS converge or diverge with that of typical development, we conducted Wilcoxon Rank-Sum tests comparing the FXS and TD groups on the six vocalization features (Table 4 and Figure 3). There were no statistically significant differences in vocalization features between the groups. Due to

significant floor effects in canonical complexity in both the FXS and TD groups, we also ran a Kruskal-Wallis Rank-Sum test to test differences between the groups on proportion of the group that used canonical syllables. This test indicated no statistically significant difference in the proportion that used canonical syllables in the FXS group (Present $n = 4$; Absent $n = 18$) versus the TD group (Present $n = 5$; Absent n = 12; $\chi^2 = .66$, $p = .415$).

Association of Vocalization Features With 24-Month Outcomes

Language Outcomes

Next, we explored whether vocalization features at 9 months were related to language outcomes at 24 months. Both groups demonstrated trends between certain vocalization features and language outcomes (Table 5; Figures 4-7). In the TD group, duration range was moderately associated with language outcomes, such that having a more narrow duration range at 9 months was associated with having higher receptive language abilities at 24 months ($\rho = -.44$, $p = .100$). In the FXS group, volubility demonstrated a trending association with language outcomes, such that infants with higher volubility at 9 months tended to have higher receptive language abilities at 24 months ($\rho = .43$, $p = .053$). Furthermore, we observed an effect of canonical syllable usage such that infants with FXS who used canonical syllables had significantly higher expressive language ability than those who did not use canonical syllables ($d = -1.05$, $p = .040$), and a trend emerged suggesting a similar effect with receptive language ability in FXS (d = -1.08, $p = .060$; Table 6). These effects were not observed in the TD group. Thus, different vocalization features demonstrated significant and trending associations with language outcomes in the FXS and TD groups, with duration range demonstrating a

stronger association with language outcomes in the TD group while volubility and canonical complexity were more strongly associated with later language abilities in FXS.

Autism Outcomes

We examined associations of vocalization features with ASD diagnosis as well as a continuous measure of ASD symptoms in the FXS group. While no vocalization features were associated with ASD diagnosis (Table 5; Figure 8), average pitch did demonstrate a trending association with ADOS-T SA CSS (Table 7; Figure 9), such that those with higher ASD symptom severity scores at 24 months tended to use vocalizations with higher pitch at 9 months (ρ = .45, p = .080). This suggests that in our sample, average pitch may be associated with ASD social affect symptoms later in development.

Exploratory Results

Due to our limited sample size and ability to detect only large, statistically-significant effects, we discuss results that demonstrate medium-sized effects that may suggest areas for further exploration in studies with larger samples. Importantly, given our small sample, these effect sizes could be spurious and thus should be further replicated and interpreted as avenues for future study.

While differences between the FXS and TD groups were not statistically significant, we observed medium-sized effects of differences between the FXS and TD group in the level of canonical complexity observed in each, the average duration and duration range of vocalizations, such that infants with FXS tended to use lower ratios of canonical syllables, longer vocalizations and a wider range of vocalization duration. These patterns are consistent with those observed in previous literature.

We also observed medium-sized effects of vocalization features with language and ASD outcomes in the hypothesized directions. In addition to associations of volubility and canonical syllable usage with language outcomes in FXS reported in our primary results, infants with higher canonical complexity at 9-months tended to have higher receptive and expressive language outcomes and those with higher average pitch tended to have lower expressive language abilities. In the TD group, in addition to associations of duration range with receptive language reported in our primary results, infants who used shorter vocalizations on average tended to have higher receptive and expressive language abilities, and those who used more narrow pitch ranges also tended to have higher receptive language outcomes. Finally, in addition to our primary results suggesting average pitch is associated with ASD symptom severity in the FXS group, infants with an ASD diagnosis tended to have longer vocalizations on average. Furthermore, canonical complexity demonstrated medium-sized effects with ASD symptom severity and ASD diagnosis, such that those with higher ASD symptom severity or an ASD diagnosis tended to use fewer canonical syllables and those who used canonical syllables tended to have higher ASD symptom severity. Overall, medium-sized effects suggest that vocalization features detect nuanced group differences and are associated with language and ASD outcomes in the expected directions.

DISCUSSION

This study presents the first comprehensive analysis of acoustic features of early vocalizations in fragile X syndrome and the association of these features with developmental outcomes in this population. As is common in FXS research, our sample size was limited, and we were underpowered to detect most small-to-medium effects; therefore, we used non-parametric analyses and set alpha at .10 for consideration of meaningful trends. Our findings supported the potential for early language features to inform later developmental and autism-related features in FXS. We report two major findings. First, the FXS and TD groups did not significantly differ in mean levels of vocalization features at 9 months, a somewhat surprising finding that may be related to aspects of the context in which vocalizations were examined. Second, several vocalization features, such as canonical complexity, volubility, and average pitch, demonstrated meaningful associations with language and ASD outcomes at 24 months in the FXS group, suggesting that these features may serve as useful early markers for delayed and atypical development in FXS. Together, these findings highlight the utility of early vocalization features as markers for later developmental outcomes in FXS.

Group Differences

We first examined whether the FXS group differed from the TD group in terms of their patterns of vocalization features. Contrary to hypotheses, infants with FXS did not demonstrate any significant differences from the TD group in mean levels of vocalization features at 9 months. These results were surprising, particularly given a recent study in which 9- to 12-month-old infants with FXS vocalized less frequently and used lower ratios of canonical syllables than age-matched TD infants (Belardi et al., 2017).

Importantly, this study examined vocalizations from retrospective home videos, in which infants were observed in more familiar contexts where their vocalization behavior may be more representative of their true ability then when they are observed in a laboratory setting interacting with an unfamiliar adult. This is consistent with previous findings suggesting that infants' vocal behavior can be influenced by the responsiveness of their caregiver (Goldstein & Schwade, 2008). Thus, the AOSI may not be an ideal scenario in which to analyze vocalizations, though further research is needed to determine the extent to which vocal behavior if affected by the setting in which it is observed and the relationship of the interaction partner to the infant.

Additionally, it is possible that a measure of volubility like overall rate may be too broad to capture nuanced differences that might be present between the two groups. In a study of vocalizations of 14-month-old TD infants and infants with ASD, overall number of vocalizations made during a 30-minute behavior sample were approximately similar between the two groups; however, group differences were identified in the number of vocalizations that were directed to another person as opposed to non-directed vocalizations (Garrido, Watson, Carballo, Garcia-Retamero, & Crais, 2017). Thus, while our findings did not detect group differences in rate of vocalizing broadly, it is possible that group differences may lie in more nuanced factors such as amount of vocalizing used for social interaction or for certain communicative functions.

The FXS and TD groups also did not significantly differ in their ratios of canonical babbling; however, it is difficult to interpret this finding given that many infants in both the TD and FXS groups did not use any canonical syllables. While previous literature suggests that canonical syllables should emerge around 6 to 10 months

of age, it is possible that at 9 months, even in typically developing samples, canonical syllable usage does not occur so frequently as to be consistently captured in 10- to 20-minute vocalization samples. Indeed, in their study of canonical syllable usage during 10-minute videos of TD and FXS infants, Belardi et al. (2017) found that only 57% of TD infants (8 of 14) and 0% of FXS infants (0 of 10) met the criterion for being the in canonical babbling stage at 9 to 12 months (i.e., ratio of canonical syllables to total speech syllables of 0.15 or greater). Importantly, canonical syllables have been identified as an early feature of infant language development that parents can easily identify and provide accurate reports about (Oller et al., 2001), making it an ideal marker of early vocalization behavior and potential index of atypical development. Thus, characterizing the emergence of canonical complexity in typical and atypical development and how it is related to later outcomes is important for determining its utility as an early marker for later developmental outcomes. Our results suggest that canonical complexity at 9 months may be associated with later language delay in FXS, but given that so few infants used any canonical syllables, it may be necessary to analyze vocalizations in longer samples or at older ages to more completely capture canonical status across a larger portion of the population.

Pitch is an important quality of prosody, which is often characterized by atypicality in individuals with ASD due to qualities like using a monotone voice or having unmodulated volume or rate of speaking. Features of vocalization pitch, such as average pitch and pitch range, have been implicated in differentiating populations with ASD from typically developing populations in older children (Fusaroli et al., 2017). Thus, we hypothesized that features of vocalization pitch in FXS may follow similar

patterns as those observed in ASD, despite our examination of vocalizations during infancy as opposed to later in development. There was no difference between the TD and FXS groups in average pitch of vocalizations, suggesting that group differences in pitch may be more specific to words and phrases as opposed to canonical and pre-canonical syllables in infancy. Our findings may also suggest that "atypical" pitch may not extend broadly to FXS, but could possibly be more specific to atypical developmental outcomes in this population, such as ASD. This is the first study of vocalization pitch in FXS; thus, further research is needed to more fully understand the role of pitch range in this population and how it compares to typical development.

Predictive Associations

We also examined associations of vocalization features with language and ASD outcomes at 24 months to determine the predictive value of vocalization features. As expected, the presence of canonical syllables was associated with higher receptive and expressive language outcomes in the FXS group. Thus, canonical syllable usage predicted language outcomes in FXS above that which can be accounted for by differing levels of developmental delay among individuals with FXS. Contrary to previous findings, the association of canonical syllable usage with language outcomes were less pronounced in the TD group. Given the wide range of receptive and expressive language scores in the TD group (Table 1), this lack of association in the TD group does not appear to be driven by limited variability of language outcomes. Thus, canonical syllable usage demonstrated a unique association with language outcomes that was specific to the FXS group. The FXS and TD groups also demonstrated unique associations of volubility and duration range, respectively, with receptive language outcomes. These findings suggest

that, similar to canonical complexity, volubility and duration range may also demonstrate distinct predictive abilities for language outcomes for the TD and FXS groups.

Finally, we examined the association of vocalization features with 24-month ASD outcomes in the FXS group. Infants with FXS who had higher pitch tended to have higher ADOS-T SA CSS scores. While no previous studies have examined the association of vocalization pitch with ASD in FXS, this is consistent with findings in idiopathic ASD populations. For example, a recent meta-analysis identified a pattern of longer and higher-pitched vocalizations in children with ASD when compared to TD children (Fusaroli et al., 2017). Importantly, this meta-analysis only included two studies of children younger than four years of age, neither of which used a sample whose age range included 9 months (0-6 months: Brisson et al., 2014; 16-48 months: Oller et al., 2010). Both of these studies reported non-significant findings regarding the pitch of vocalizations for children with ASD compared to TD children, suggesting that the association of pitch with ASD may vary by age. Given the high prevalence of ASD in FXS, our finding suggesting that average pitch may predict ASD risk has important implications on improving the early identification of ASD in this population. Furthermore, our findings in FXS correspond to those reported in studies with other high-risk populations; thus, we build on evidence that some early features of ASD may broadly predict ASD risk across a variety of risk groups. This information can in turn inform etiological pathways, particularly when risk groups with genetically-identified causes of ASD are included, as is the case with FXS. Overall, our findings indicating associations of average pitch with ASD outcomes in FXS have important implications for early identification of ASD in this population, and further research should continue to

characterize this features early in development to further inform our understanding of this features and its association with ASD.

Summary

This study characterizes early vocalization features in FXS, demonstrating how these features are associated with language ability and ASD outcomes, as well as highlighting how these features in FXS may diverge from patterns observed in typically developing populations. Results suggested that while mean levels of vocalization features may not differ between TD and FXS, within-group variability in FXS may predict later language outcomes. Furthermore, average pitch demonstrated a meaningful association with later ASD outcomes in the FXS group, emphasizing the potential of this features as an early marker of ASD in this population. Importantly, characterizing the associations of early vocalization features with ASD outcomes in FXS can inform mechanisms of ASD development that can then be tested broadly with other etiologically-distinct populations at risk for ASD.

Limitations and Future Directions

This study presents one of the few studies that has examined vocalization features in young infants with FXS, and the first study to examine acoustic properties of infant vocalizations in FXS. Despite these strengths, this study does pose some limitations that are important to consider when interpreting the results and planning future research in this area. First, the small sample used in this study limited our power to detect nuanced effects and increased the likelihood of attaining spurious effects, particularly when small or medium effects were indicated. Thus, the results of this study should be viewed as preliminary data that future studies with larger samples can use to directly test these

hypothesized effects. Secondly, as previously mentioned, it is possible that the AOSI may not be an ideal context in which to analyze infant vocalizations, as this is a relatively short sample of vocal behavior and there is a chance that infants vocalize less frequently due to the presentation of an assortment of novel and potentially distracting toys or when interacting with an unfamiliar adult. Thus, it will be important for future studies to examine features of infants' vocalizations in more naturalistic settings and across longer periods of vocalization and with different interaction partners. Finally, the cross-sectional nature of this study precluded us from characterizing the developmental trajectories of these vocalization features and any factors that may influence these trajectories. Therefore, future research should analyze vocalization features longitudinally to more fully characterize how these features are related to development and outcomes in FXS.

As stated above, examining vocalization features in more naturalistic settings and using longer samples of vocalization behavior is an important future direction for the study of early vocalization features in FXS. The Language Environment Analysis (LENA) system, which collects day-long samples of the child's vocalization behavior in their home and community using a small, automated audio recorder, offers one possible method for extending the study of early vocalization features to more representative samples of vocalization behavior in FXS. Additionally, while this study provides critical information about how vocalization features in FXS compare with those of typical development, it will be important for future studies to include comparisons with other groups with atypical development, including populations with other neurogenetic syndromes or those at increased risk for developmental delay. Comparisons with other high-risk populations can provide insight into the robustness of the associations of

vocalization features with typical and atypical development, which in turn can inform possible mechanisms of etiology and improve early identification of early childhood disorders, such as ASD.

Conclusion

This study presents the first comprehensive analysis of early vocalization features in FXS. We present preliminary evidence that features of early vocalizations in FXS may not differ meaningfully from those observed in typical development when examined during a short, semi-structured interaction, though these groups may differ in the early vocalization features that are most relevant to later language outcomes. We also identified a vocalization feature, average pitch, that demonstrated an association with later ASD outcome in FXS, suggesting that this feature may be a potential early marker for ASD in this population. Because FXS is characterized by such high rates of ASD which have a known single-gene cause, identifying early markers for ASD in this population may inform further research on mechanisms through which ASD develops in this population and in other high-risk populations which have no known etiological mechanisms for ASD. Thus, further research and deeper understanding of these early vocalization features in typical and atypical development may lead to improved early identification methods, treatment approaches, and overall well-being of individuals in the ASD population.

LIST OF REFERENCES

Abbeduto, L., Brady, N., & Kover, S. T. (2007). Langauge development and fragile X syndrome: Profiles, syndrome-specificity, and within-syndrome differences. *Mental Retardation and Developmental Disabilities Research Reviews*, *13*, 36-46. https://doi.org/10.1002/mrdd.20142

Abbeduto, L., McDuffie, A., & Thurman, A. J. (2014). The fragile x syndrome-autism comorbidity: What do we really know? *Frontiers in Genetics*, 5, 355. https://doi.org/10.3389/fgene.2014.00355

Belardi, K., Watson, L. R., Faldowski, R. A., Hazlett, H., Crais, E., Baranek, G. T., . . . Oller, D. K. (2017). A retrospective video analysis of canonical babbling and volubility in infants with fragile X syndrome at 9-12 months of age. *Journal of Autism and Developmental Disorders*, *47*(4), 1193-1206. https://doi.org/10.1007/s10803-017-3033-4

Bhakar, A. L., Dölen, G., & Bear, M. F. (2012). The pathophysiology of fragile X (and what it teaches us about synapses). *Annual Review of Neuroscience*, *35*, 417-443. https://doi.org/10.1146/annurev-neuro-060909-153138

Boersma, P., & Weenink, D. (2016). *Praat: doing phonetics by computer*. Retrieved from http://www.praat.org/

Brady, N., Skinner, D., Roberts, J., & Hennon, E. (2006). Communication in young children with fragile X syndrome: A qualitative study of mothers' perspectives. *American Journal of Speech-Language Pathology*, *15*(4), 353-364. https://doi.org/10.1044/1058-0360(2006/033)

Brisson, J., Martel, K., Serres, J., Sirois, S., & Adrien, J. L. (2014). Acoustic analysis of oral productions of infants later diagnosed with autism and their mother. *Infant Mental Health Journal, 35*(3), 285-295. https://doi.org/10.1002/imhj.21442

Bryson, S. E., Zwaigenbaum, L., McDermott, C., Rombough, V., & Brian, J. (2008). The autism observation scale for infants: Scale development and reliability data. *Journal of Autism and Developmental Disorders, 38*(4), 731-738. https://doi.org/10.1007/s10803-007-0440-y

Budimirovic, D. B., Bukelis, I., Cox, C., Gray, R. M., Tierney, E., & Kaufmann, W. E. (2006). Autism spectrum disorder in fragile X syndrome: Differential contribution of adaptive socialization and social withdrawal. *American Journal of Medical Genetics, Part A, 140*(17), 1814-1826. https://doi.org/10.1002/ajmg.a.31405

Budimirovic, D. B., & Kaufmann, W. E. (2011). What can we learn about autism from studying fragile X syndrome? *Developmental Neuroscience, 33*(5), 379-394. https://doi.org/10.1159/000330213

Diagnostic and statistical manual of mental disorders (DSM-5®). (2013). Washington, DC: American Psychiatric Publication.

Esler, A. N., Bal, V. H., Guthrie, W., Wetherby, A., Weismer, S. E., & Lord, C. (2015). The Autism Diagnostic Observation Schedule, Toddler Module: Standardized severity scores. *Journal of Autism and Developmental Disorders, 45*(9), 2704-2720. https://doi.org/10.1007/s10803-015-2432-7

Filipe, M. G., Frota, S., Castro, S. L., & Vicente, S. G. (2014). Atypical prosody in Asperger syndrome: Perceptual and acoustic measurements. *Journal of Autism and Developmental Disorders*, *44*(8), 1972-1981. https://doi.org/10.1007/s10803-014-2073-2

Finestack, L. H., Richmond, E. K., & Abbeduto, L. (2009). Language development in individuals with fragile X syndrome. *Topics in Language Disorders*, *29*(2), 133-148. https://doi.org/10.1097/TLD.0b013e3181a72016

Franklin, B., Warlaumont, A. S., Messinger, D., Bene, E., Nathani Iyer, S., Lee, C.-C., . . . Oller, D. K. (2013). Effects of parental interaction on infant vocalization rate, variability and vocal type. *Language Learning and Development*, *10*(3), 279-296. https://doi.org/10.1080/15475441.2013.849176

Fusaroli, R., Lambrechts, A., Bang, D., Bowler, D. M., & Gaigg, S. B. (2017). Is voice a marker for Autism spectrum disorder? A systematic review and meta-analysis. *Autism Research*, *10*(3), 384-407. https://doi.org/10.1002/aur.1678

Garrido, D., Watson, L. R., Carballo, G., Garcia-Retamero, R., & Crais, E. R. (2017). Infants at-risk for autism spectrum disorder: Patterns of vocalizations at 14 months. *Autism Research*, *10*(8), 1372-1383. https://doi.org/10.1002/aur.1788

Goldstein, M. H., & Schwade, J. a. (2008). Social feedback to babbling facilitates vocallLearning. *Psychological Science*, *19*(5), 515-523. https://doi.org/10.1111/j.1467-9280.2008.02117.x

Goldstein, M. H., Schwade, J. A., & Bornstein, M. H. (2009). The value of vocalizing: Five-month-old infants associate their own noncry vocalizations with responses from caregivers. *Child Development*, *80*(3), 636-644. https://doi.org/10.1111/j.1467-8624.2009.01287.x

Gratier, M., & Devouche, E. (2011). Imitation and repetition of prosodic contour in vocal interaction at 3 months. *Developmental Psychology*, *47*(1), 67-76. https://doi.org/10.1037/a0020722

Hall, S. S., Lightbody, A. A., Hirt, M., Rezvani, A., & Reiss, A. L. (2010). Autism in fragile X syndrome: A category mistake? *Journal of the American Academy of Child & Adolescent Psychiatry*, *49*(9), 921-933. https://doi.org/10.1016/j.jaac.2010.07.001.Autism

Hall, S. S., Lightbody, A. A., & Reiss, A. L. (2008). Compulsive, self-injurious, and autistic behavior in children and adolescents with fragile X syndrome. *American Journal on Mental Retardation*, *113*(1), 44-53. https://doi.org/10.1352/0895-8017(2008)113[44:CSAABI]2.0.CO;2

Harris, S. W., Hessl, D., Goodlin-Jones, B., Ferranti, J., Bacalman, S., Barbato, I., . . . Hagerman, R. J. (2008). Autism profiles of males with fragile X syndrome. *American Journal on Mental Retardation*, *113*(6), 427-438. https://doi.org/10.1352/2008.113:427-438

Iyer, S. N., & Oller, D. K. (2008). Prelinguistic vocal development in infants with typical hearing and infants with severe-to-profound hearing loss. *Volta Review*, *108*(2), 115-138. https://doi.org/10.1002/ana.22528.Toll-like

Kau, A. S. M., Tierney, E., Bukelis, I., Stump, M. H., Kates, W. R., Trescher, W. H., & Kaufmann, W. E. (2004). Social behavior profile in young males with fragile X syndrome: Characteristics and specificity. *American Journal of Medical Genetics, 126A*(1), 9-17. https://doi.org/10.1002/ajmg.a.20218

Kaufmann, W. E., Cortell, R., Kau, A. S. M., Bukelis, I., Tierney, E., Gray, R. M., . . . Stanard, P. (2004). Autism spectrum disorder in fragile X syndrome: communication, social interaction, and specific behaviors. *American Journal of Medical Genetics Part A, 129*(3), 225-234.

Klusek, J., Martin, G. E., & Losh, M. (2014). A comparison of pragmatic language in boys with autism and fragile X syndrome. *Journal of Speech, Language, and Hearing Research, 57*(5), 1692-1707. https://doi.org/10.1044/2014_JSLHR-L-13-0064

Le Couteur, A., Lord, C., & Rutter, M. (2003). *The autism diagnostic interview-revised (ADI-R)*. Los Angeles, CA: Western Psychological Services.

Levy, Y., Gottesman, R., Borochowitz, Z., Frydman, M., & Sagi, M. (2006). Language in boys with fragile X syndrome. *Journal of Child Language, 33*(01), 125-144. https://doi.org/10.1017/S030500090500718X

Lord, C. (2012). A multisite study of the clinical diagnosis of different autism spectrum disorders. *Archives of General Psychiatry, 69*(3), 306-316. https://doi.org/10.1001/archgenpsychiatry.2011.148

Lord, C., Rutter, M., DiLavore, P., Risi, S., Gotham, K., & Bishop, S. (2012). *Autism Diagnostic Observation Schedule–2nd edition (ADOS-2)*. Los Angeles, CA: Western Psychological Corporation.

Losh, M., Martin, G. E., Klusek, J., Hogan-Brown, A. L., & Sideris, J. (2012). Social communication and theory of mind in boys with autism and fragile X syndrome. *Frontiers in Psychology, 3*(August), 1-12. https://doi.org/10.3389/fpsyg.2012.00266

Lynch, M. P., Oller, D. K., Steffens, M. L., & Buder, E. H. (1995). Phrasing in prelinguistic vocalizations. *Developmental Psychobiology, 28*(1), 3-25. https://doi.org/10.1002/dev.420280103

Lynch, M. P., Oller, D. K., Steffens, M. L., & Levine, S. L. (1995). Onset of speech-like vocalizations in infants with Down syndrome. *American Journal on Mental Retardation, 100*, 68-86.

Marcos, H. (1987). Communicative functions of pitch range and pitch direction in infants. *Journal of Child Language, 14*(2), 255-268. https://doi.org/10.1017/S0305000900012915

Marschik, P. B., Bartl-Pokorny, K. D., Sigafoos, J., Urlesberger, L., Pokorny, F., Didden, R., . . . Kaufmann, W. E. (2014). Development of socio-communicative skills in 9- to 12-month-old individuals with fragile X syndrome. *Research in Developmental Disabilities, 35*(3), 597-602. https://doi.org/10.1016/j.ridd.2014.01.004

Martin, G. E., Losh, M., Estigarribia, B., Sideris, J., & Roberts, J. (2013). Longitudinal profiles of expressive vocabulary, syntax and pragmatic language in boys with fragile X syndrome or Down syndrome. *International Journal of Language & Communication Disorders, 48*(4), 432-443. https://doi.org/10.1111/1460-6984.12019

McCary, L., & Roberts, J. (2013). Early identification of autism in fragile X syndrome: a review. *Journal of Intellectual Disability*, *57*(9), 803-814. https://doi.org/10.1111/j.1365-2788.2012.01609.x.Early

McDuffie, A., Thurman, A. J., Hagerman, R. J., & Abbeduto, L. (2015). Symptoms of autism in males with fragile X syndrome: A comparison to nonsyndromic ASD using current ADI-R scores. *Journal of Autism and Developmental Disorders, 45*(7), 1925-1937. https://doi.org/10.1007/s10803-013-2013-6

McGillion, M., Herbert, J. S., Pine, J., Vihman, M., DePaolis, R., Keren-Portnoy, T., & Matthews, D. (2017). What paves the way to conventional language? The predictive value of babble, pointing, and socioeconomic Status. *Child Development*, *88*(1), 156-166. https://doi.org/10.1111/cdev.12671

Morgan, L., & Wren, Y. E. (2018). A systematic review of the literature on early vocalizations and babbling patterns in young children. *Communication Disorders Quarterly*. https://doi.org/10.1177/1525740118760215

Mullen, E. M. (1995). *Mullen scales of early learning*. Circle Pines, MN: American Guidance Service.

Munson, J., Dawson, G., Sterling, L., Beauchaine, T., Zhou, A., Koehler, E., . . . Abbott, R. (2008). Evidence for latent classes of IQ in young children with autism spectrum disorder. *American Journal on Mental Retardation*, *113*(6), 439-452. https://doi.org/10.1352/2008.113:439-452

Oller, D. K., & Eilers, R. E. (1988). The role of audition in infant Babbling. *Child Development*, *59*(2), 441-449. https://doi.org/10.2307/1130323

Oller, D. K., Eilers, R. E., & Basinger, D. (2001). Intuitive identification of infant vocal sounds by parents. *Developmental Science, 4*(1), 49-60. https://doi.org/10.1111/1467-7687.00148

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*(4), 223-245. https://doi.org/10.1016/S0021-9924(99)00013-1

Oller, D. K., Eilers, R. E., Neal, A. R., & Schwartz, H. K. (2006). Precursors to speech in infancy: The prediction of speech and language disorders. *Journal of Communication Disorders, 32*(207), 223-245. https://doi.org/http://dx.doi.org/10.1016/S0021-9924(99)00013-1

Oller, D. K., Eilers, R. E., Steffens, M. L., Lynch, M. P., & Urbano, R. (1994). Speech-like vocalizations in infancy: An evaluation of potential risk factors. *Journal of Child Language*, *21*(01), 33-58. https://doi.org/10.1017/S0305000900008667

Oller, D. K., Niyogi, P., Gray, S., Richards, J. a, Gilkerson, J., Xu, D., . . . Warren, S. F. (2010). Automated vocal analysis of naturalistic recordings from children with autism, language delay, and typical development. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(30), 13354-13359. https://doi.org/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*(10), 2413-2428. https://doi.org/10.1007/s10803-014-2047-4

Paul, R., Fuerst, Y., Ramsay, G., Chawarska, K., & Klin, A. (2011). Out of the mouths of babes: Vocal production in infant siblings of children with ASD. *Journal of Child Psychology and Psychiatry, 52*(5), 588–598. https://doi.org/10.1111/j.1469-7610.2010.02332.x

Pavtube Video Converter. (2016). Retrieved from http://www.pavtube.com

Plumb, A. M., & Wetherby, A. M. (2013). Vocalization development in toddlers with autism spectrum disorder. *Journal of Speech, Language, and Hearing Research, 56*, 721-734. https://doi.org/10.1044/1092-4388)

Quigley, J., McNally, S., & Lawson, S. (2016). Prosodic patterns in interaction of low-risk and at-risk-of-autism spectrum disorders infants and their mothers at 12 and 18 months. *Language Learning and Development*, *5441*(June), 1-16. https://doi.org/10.1080/15475441.2015.1075405

Rague, L., Caravella, K., Tonnsen, B., Klusek, J., & Roberts, J. E. (2018). Early gesture use in fragile X syndrome. *Journal of Intellectual Disability Research, 62*(7), 625-636. https://doi.org/10.1111/jir.12498

Roberts, J. E., Mirrett, P., Anderson, K., Burchinal, M., & Neebe, E. (2002). Early communication, symbolic behavior, and social profiles of young males with fragile X syndrome. *American Journal of Speech-Language Pathology, 11*, 295-304. https://doi.org/10.1044/1058-0360(2002/034)

Roberts, J. E., Tonnsen, B. L., McCary, L. M., Caravella, K. E., & Shinkareva, S. V. (2016). Brief report: Autism symptoms in infants with fragile X syndrome. *Journal of Autism and Developmental Disorders*, *46*(12), 3830-3837. https://doi.org/10.1007/s10803-016-2903-5

Roberts, J. E., Weisenfeld, L. A. H., Hatton, D. D., Heath, M., & Kaufmann, W. E. (2007). Social approach and autistic behavior in children with fragile X syndrome. *Journal of Autism and Developmental Disorders*, *37*(9), 1748-1760. https://doi.org/10.1007/s10803-006-0305-9

Roy, S., Watkins, N., & Heck, D. (2012). Comprehensive analysis of ultrasonic vocalizations in a mouse model of fragile X syndrome reveals limited , call type specific deficits, *7*(9), 1-6. https://doi.org/10.1371/journal.pone.0044816

Santoro, M. R., Bray, S. M., & Warren, S. T. (2012). Molecular mechanisms of fragile X syndrome: A twenty-year perspective. *Annual Review of Pathology: Mechanisms of Disease*, *7*, 219-245. https://doi.org/10.1146/annurev-pathol-011811-132457

Schoen, E., Paul, R., & Chawarska, K. (2012). Phonology & vocal behavior in toddlers with ASDs. *Autism Research*, *4*(3), 177-188. https://doi.org/10.1002/aur.183.Phonology

Sharda, M., Subhadra, T. P., Sahay, S., Nagaraja, C., Singh, L., Mishra, R., . . . Singh, N. C. (2010). Sounds of melody-pitch patterns of speech in autism. *Neuroscience Letters*, *478*(1), 42-45. https://doi.org/10.1016/j.neulet.2010.04.066

Sheinkopf, S. J., Iverson, J. M., Rinaldi, M. L., & Lester, B. M. (2012). Atypical cry acoustics in 6-month-old infants at risk for autism spectrum disorders. *Autism Research*, *5*(5), 331-339. https://doi.org/10.1002/aur.1244

Snow, D., & Balog, H. L. (2002). Do children produce the melody before the words? A review of developmental intonation research. *Lingua*, *112*(12), 1025-1058. https://doi.org/10.1016/S0024-3841(02)00060-8

Vihman, M., & Vihman, V.-A. (2011). From first words to segments: A case study in phonological development. *Experience, Variation and Generalization: Learning a First Language*, *7*, 109-134. https://doi.org/10.1075/tilar.7.07vih

Warren, S. F., Gilkerson, J., Richards, J. A., Oller, D. K., Xu, D., Yapanel, U., & Gray, S. (2010). What automated vocal analysis reveals about the vocal production and language learning environment of young children with autism. *Journal of Autism and Developmental Disorders*, *40*, 555-569. https://doi.org/10.1007/s10803-009-0902-5

Wilson, D., & Wharton, T. (2006). Relevance and prosody. *Journal of Pragmatics*, *38*(10), 1559-1579. https://doi.org/10.1016/j.pragma.2005.04.012

Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., & Sloetjes, H. (2006). ELAN: A professional framework for multimodality research. *Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC)*, 1556-1559. https://doi.org/10.3758/BRM.41.3.591

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*(5), 1638-1650. https://doi.org/10.1044/2014_JSLHR-S-13-0037

Zwaigenbaum, L., Bauman, M. L., Stone, W. L., Yirmiya, N., Estes, A., Hansen, R. L., . . . Wetherby, A. M. (2015). Early identification of autism spectrum disorder: Recommendations for practice and research. *Pediatrics*, *136*, S10-S40. https://doi.org/10.1542/peds.2014-3667C

APPENDIX A

Table 1

Effect Sizes (d) for Standardized Mean Differences and Variance Ratios (var) of Core Variables for Matching Comparison Groups

	FXS (*n* = 22)		TD (*n* = 17)			
FXS vs. TD	*M (SD)*	Range	*M (SD)*	Range	*d*	var
9-Month CA	9.49 (0.76)	7.69-10.85	9.31 (0.44)	8.58-0.46	.29	3.07
9-Month NVMA	7.86 (2.83)	2.5-13.0	9.56 (1.71)	5.0-12.5	.70	2.73
24-Month CA	25.13 (1.60)	22.99-29.62	24.70 (0.75)	23.57-26.32	.33	4.55
24-Month NVMA	18.05 (4.51)	7.5-26.0	23.24 (1.96)	19.5-27.5	1.43	5.28
24-Month DD Level[a]	7.09 (4.61)	-1.24-18.97	1.47 (2.33)	-3.29-6.41	1.48	3.91
24-Month MSEL-RL AE	14.45 (6.25)	4-27	27.59 (3.83)	20-37	2.46	2.67
24-Month MSEL-EL AE	14.59 (6.53)	3-29	22.53 (4.61)	14-29	1.37	2.00

(*table continues*)

FXS-O vs. FXS-ASD	FXS-O (*n* = 12)		FXS-ASD (*n* = 10)			
	M (SD)	Range	*M (SD)*	Range	*d*	var
9-Month CA	9.46 (0.67)	8.22-10.42	9.54 (0.90)	7.69-10.85	.10	1.85
9-Month NVMA	8.13 (3.41)	2.5-13.0	7.55 (2.07)	3.5-10.5	.20	2.74
24-Month CA	25.41 (1.88)	22.99-29.62	24.80 (1.20)	23.32-26.70	.38	2.48
24-Month NVMA	19.83 (3.41)	16.0-26.0	15.90 (4.88)	7.5-23.5	.95	2.04
24-Month DD Level[a]	5.58 (3.63)	-1.24-10.62	8.90 (5.17)	1.14-18.97	.76	2.03
24-Month MSEL-RL AE	17.17 (5.95)	7-27	11.20 (5.14)	4-20	1.07	1.34
24-Month MSEL-EL AE	16.92 (6.23)	5-29	11.80 (6.01)	3-22	.83	1.07
24-Month ADOS-T SA CSS	2.33 (1.72)	1-7	7.00 (2.05)	3-9	2.48	1.42

Note. CA = chronological age; NVMA = nonverbal mental age; DD Level = developmental delay level; MSEL-RL AE = Mullen receptive language age equivalent; MSEL-EL AE = Mullen expressive language age equivalent; ADOS-T SA CSS = ADOS-T Social Affect Calibrated Severity Score. [a]24-Month Developmental Delay is calculated by subtracting the child's mental age (i.e., the average of their MSEL Visual Reception and Fine Motor age equivalents) from their chronological age in months. As per thresholds reported in Kover & Atwood (2013), groups with a standardized mean difference with a Cohen's *d* of less than .20 and a variance ratio less than 1.33 are considered adequately matched.

Table 2

Demographic Information

	FXS-O (*n* = 12)	FXS-ASD (*n* = 10)	TD (*n* = 17)
Race/Ethnicity			
Race			
White	7 (58%)	7 (70%)	14 (82%)
Black or African American	2 (17%)	0 (0%)	2 (12%)
More than One Race	3 (25%)	3 (30%)	1 (6%)
Not Reported	0 (0%)	0 (0%)	0 (0%)
Ethnicity			
Hispanic/Latino	2 (17%)	0 (0%)	1 (6%)
Not Hispanic/Latino	10 (83%)	10 (100%)	16 (94%)
Household Income			
$0 - $15,000	1 (8%)	1 (10%)	1 (6%)
$15,001 - $35,000	1 (8%)	0 (0%)	2 (12%)
$35,001 - $75,000	4 (33%)	3 (30%)	5 (29%)
$75,001 - $150,000	1 (8%)	3 (30%)	6 (35%)
Over $150,000	0 (0%)	1 (10%)	0 (0%)
Not Reported	5 (42%)	2 (20%)	3 (18%)

(*table continues*)

	FXS-O (n = 12)	FXS-ASD (n = 10)	TD (n = 17)
Maternal Education Level			
Less Than High School	1 (8%)	1 (10%)	0 (0%)
High School Degree	1 (8%)	0 (0%)	1 (6%)
Associates Degree	1 (8%)	1 (10%)	2 (12%)
Some College	1 (8%)	2 (20%)	2 (12%)
Bachelor's Degree	6 (50%)	2 (20%)	6 (35%)
More Than Bachelor's Degree	2 (17%)	4 (40%)	6 (35%)
Not Reported	0 (0%)	0 (0%)	0 (0%)

Table 3

Rate and Number of Canonical Syllables Used Among Canonical Syllable Users in the FXS and TD Groups

	FXS (n = 4)		TD (n = 5)	
	M	Range	*M*	Range
Canonical Syllables	2.75	1-5	3	1-8
Complexity Ratio	10.33	5.6-12.8	15.04	4.3-34.8

Note. Canonical ratio = number of canonical syllables divided by number of total speech syllables.

Table 4

Wilcoxon Rank-Sum Tests of Differences in 9-Month Vocalization Features Between FXS and TD

	FXS (n = 12)		TD (n = 10)				
	M	*SD*	*M*	*SD*	*W*	*p*	*d*
Volubility	1.25	1.11	1.25	0.91	180	.845	.00
Canonical Complexity	1.88	4.25	4.42	9.32	164.5	.399	-.37
Average Duration	1.39	0.92	1.02	0.96	241	.131	.39
Duration Range[a]	5.27	4.42	2.61	2.68	206	.149	.69
Average Pitch[b]	337.29	55.06	334.48	50.62	129	.970	.05
Pitch Range[b]	157.59	100.40	176.15	105.70	98.5	.630	-.18

Note. $^{\dagger}p < .10$, $*p < .05$, $**p < .01$, $***p < .001$. [a]Samples used for correlations involving duration range are smaller (FXS n = 20; TD n = 16) due to some participants having too few eligible vocalizations (<2) to calculate a range. [b]Samples used for correlations involving average pitch and pitch range are smaller (FXS n = 17; TD n = 13) due to some participants having 1 or fewer vocalizations with Codability = 1.

Table 5

Spearman Rank-Order Correlations of 9-Month Vocalization Features With 24-Month Outcomes

	Controlling for 24-Month Level of Developmental Delay					
	24M MSEL-RL AE		24M MSEL-EL AE		24M ADOS-T SA CSS	
FXS (n = 22)	ρ	p	ρ	p	ρ	p
Volubility	.43†	.053	.13	.549	-.10	.652
Canonical Complexity	.32	.163	.35	.115	-.13	.563
Average Duration	.24	.287	.23	.322	.11	.644
Duration Range[a]	.29	.228	.22	.359	.06	.813
Average Pitch[b]	-.13	.618	-.38	.147	.45†	.080
Pitch Range[b]	.15	.576	-.20	.451	-.19	.493

	Controlling for 24-Month Level of Developmental Delay			
	24M MSEL-RL AE		24M MSEL-EL AE	
TD (n = 17)	ρ	p	ρ	p
Volubility	-.11	.698	.04	.883
Canonical Complexity	-.04	.883	-.08	.775
Average Duration	-.31	.246	-.30	.268
Duration Range[a]	-.44†	.100	-.03	.908
Average Pitch[b]	.12	.700	.26	.421
Pitch Range[b]	-.35	.262	-.10	.769

Note. †p < .10, *p < .05, **p < .01, ***p < .001. Correlations of vocalization features with ADOS-T SA CSS scores were not calculated for the TD group given the inclusion criteria for this group precluded having a severity score of greater than 3. [a]Samples used for correlations involving duration range are smaller (FXS n = 20; TD n = 16) due to some participants having too few eligible vocalizations (<2) to calculate a range. [b]Samples used for correlations involving average pitch and pitch range are smaller (FXS n = 17; TD n = 13) due to some participants having 1 or fewer vocalizations with Codability = 1.

Table 6

Wilcoxon Rank-Sum Tests of Differences in 24-Month Outcomes Between Participants With Canonical Syllables Present vs. Absent at 9-Months

	Absent (n = 18)		Present (n = 4)				
FXS (n = 22)	*M*	*SD*	*M*	*SD*	*d*	*W*	*p*
24-Month MSEL-RL AE	13.28	6.06	19.75	4.50	-1.08	13.5	.060†
24-Month MSEL-EL AE	13.39	6.35	20.00	4.69	-1.05	11.5	.040*
24-Month ADOS SA CSS	4.72	2.80	3.25	4.03	.48	47.5	.343
	Absent (n = 12)		Present (n = 5)				
TD (n = 17)	*M*	*SD*	*M*	*SD*	*d*	*W*	*p*
24-Month MSEL-RL AE	27.27	3.49	27.20	4.97	.14	30	1.00
24-Month MSEL-EL AE	22.33	4.62	23.00	5.10	-.14	29	.958

Note. $^{\dagger}p < .10$, $*p < .05$, $**p < .01$, $***p < .001$.

Table 7

Wilcoxon Rank-Sum Tests of Differences in 9-Month Vocalization Features Between FXS Participants With and Without ASD at 24 Months

	FXS-O (n = 12)		FXS-ASD (n = 10)				
	M	*SD*	*M*	*SD*	*d*	*W*	*p*
Volubility	1.20	1.08	1.31	1.21	.09	56	.817
Canonical Complexity	2.52	4.85	1.11	3.51	-.32	50.5	.378
Average Duration	1.20	0.84	1.62	1.00	.44	74	.381
Duration Range[a]	5.09	4.81	5.49	4.17	.09	53	.824
Average Pitch[b]	334.15	63.25	340.82	48.28	.12	40	.743
Pitch Range[b]	167.00	126.79	147.00	66.52	-.19	34	.888

Note. $^{\dagger}p < .10$, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$. [a]Samples used for correlations involving duration range are smaller (FXS n = 20; TD n = 16) due to some participants having too few eligible vocalizations (<2) to calculate a range. [b]Samples used for correlations involving average pitch and pitch range are smaller (FXS n = 17; TD n = 13) due to some participants having 1 or fewer vocalizations with Codability = 1.

APPENDIX B

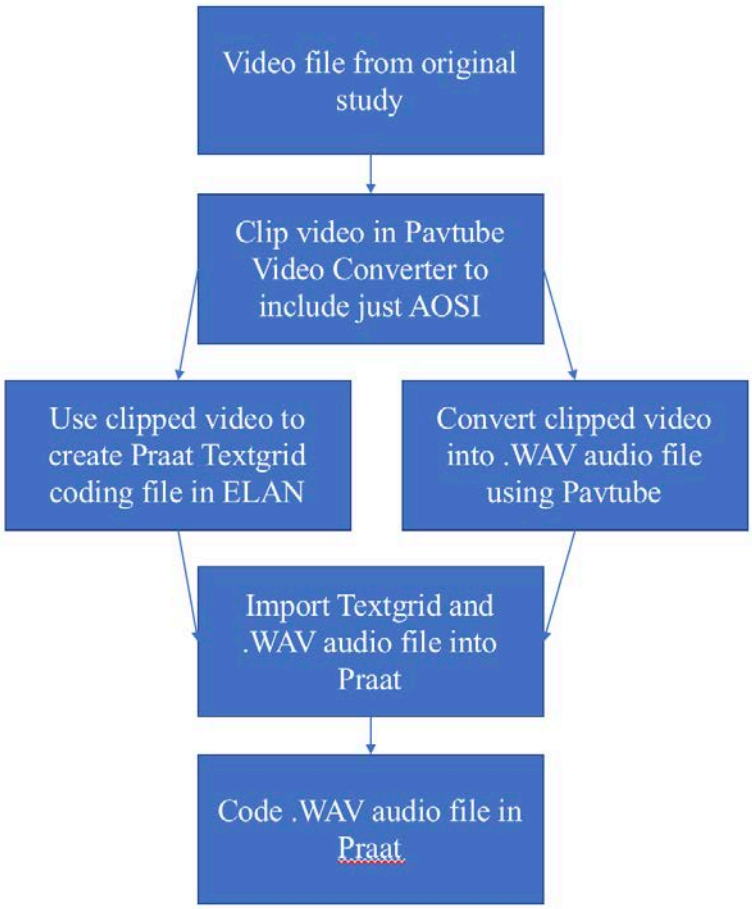
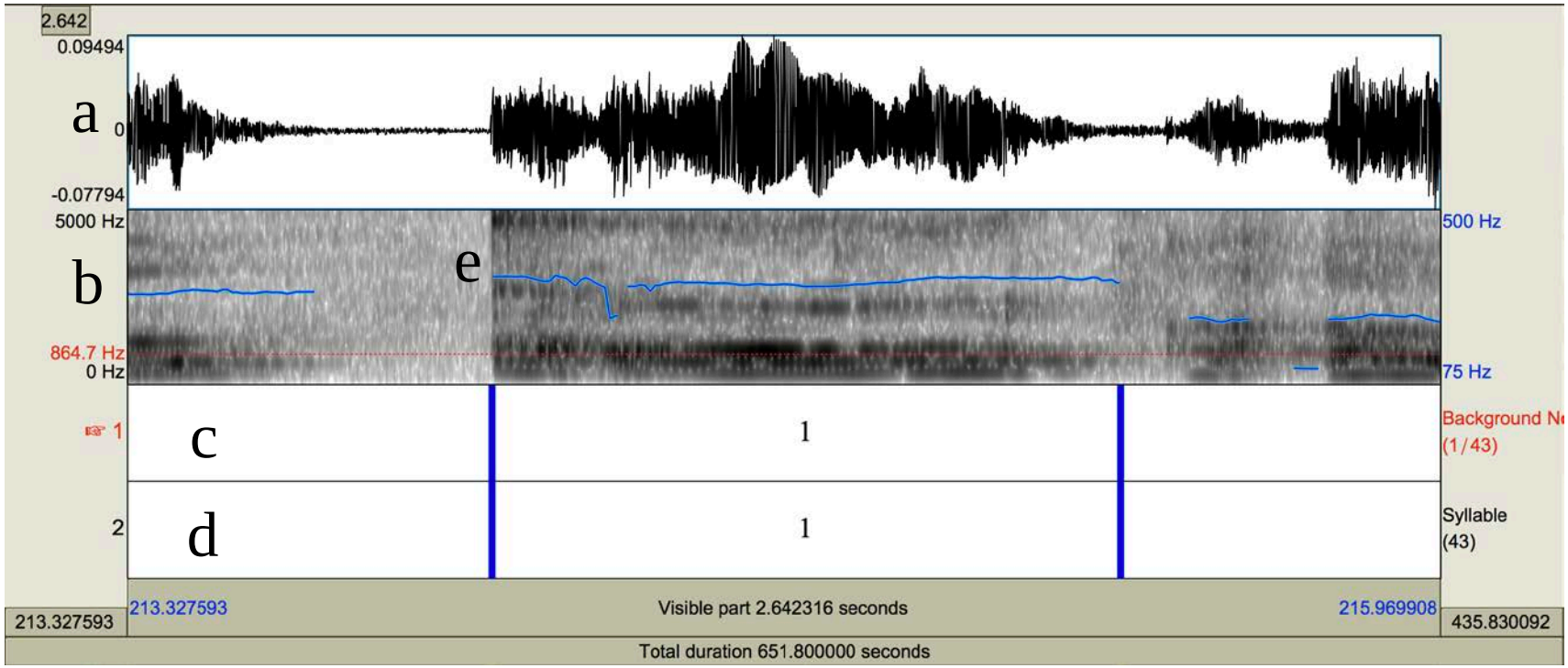


Figure 1. Process for preparation of video files for coding.



Note. (a) shows the waveform, (b) shows the spectogram, and text tiers in (c) and (d) are used for coding noise and syllables respectively. In this image, a single vocalization has been coded as Pre-canonical (Speech Type = 1) and as having no background noise (Codability = 1). (e) shows the pitch track.

Figure 2. Image of Praat coding window.

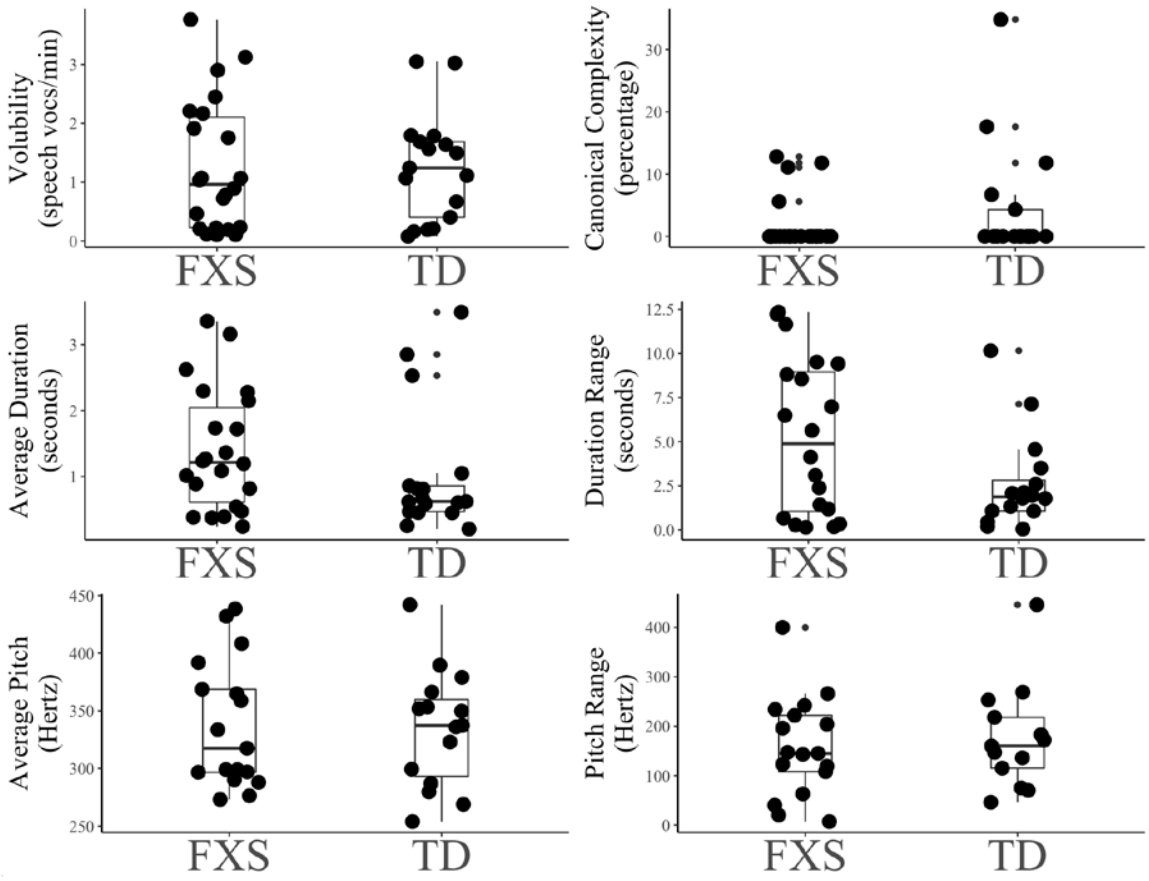


Figure 3. Comparison of vocalization features between FXS and TD participants.

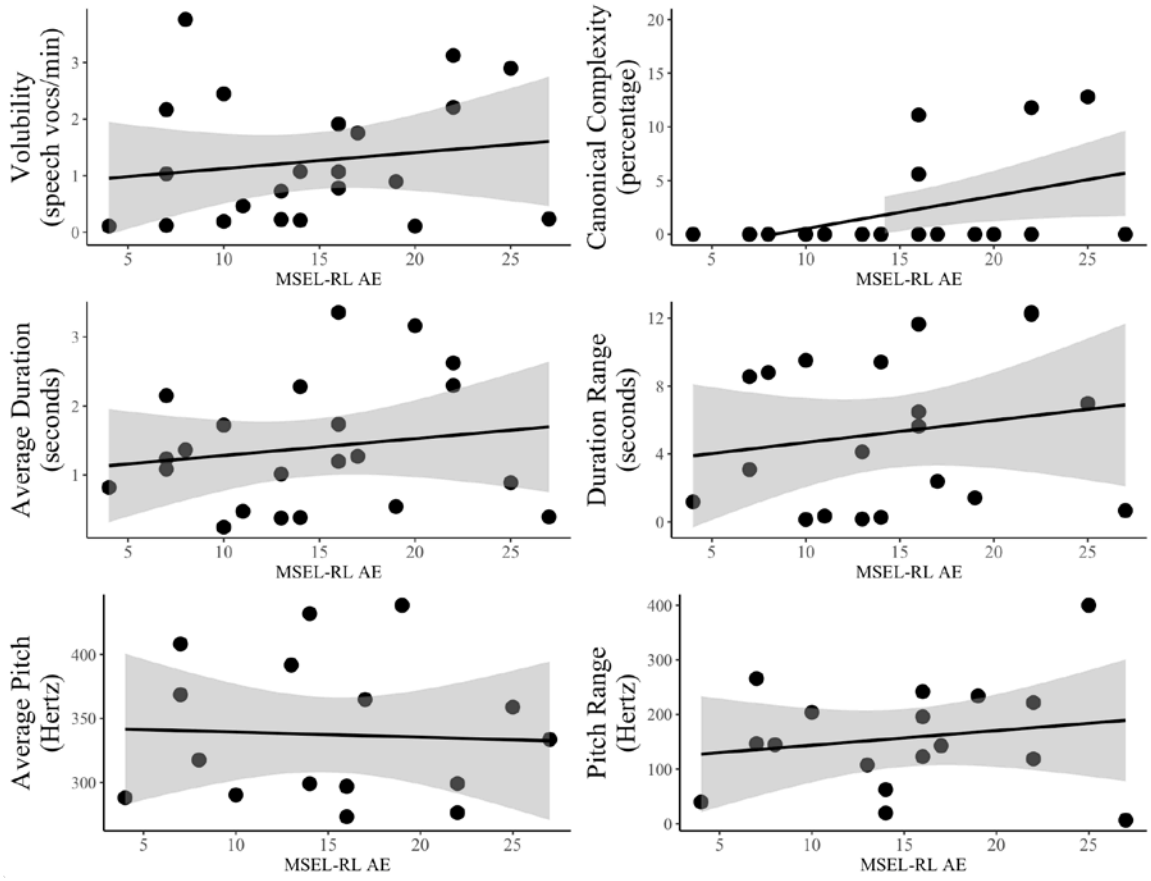


Figure 4. Associations of vocalization features with 24-month Mullen Receptive Language age equivalents in FXS.

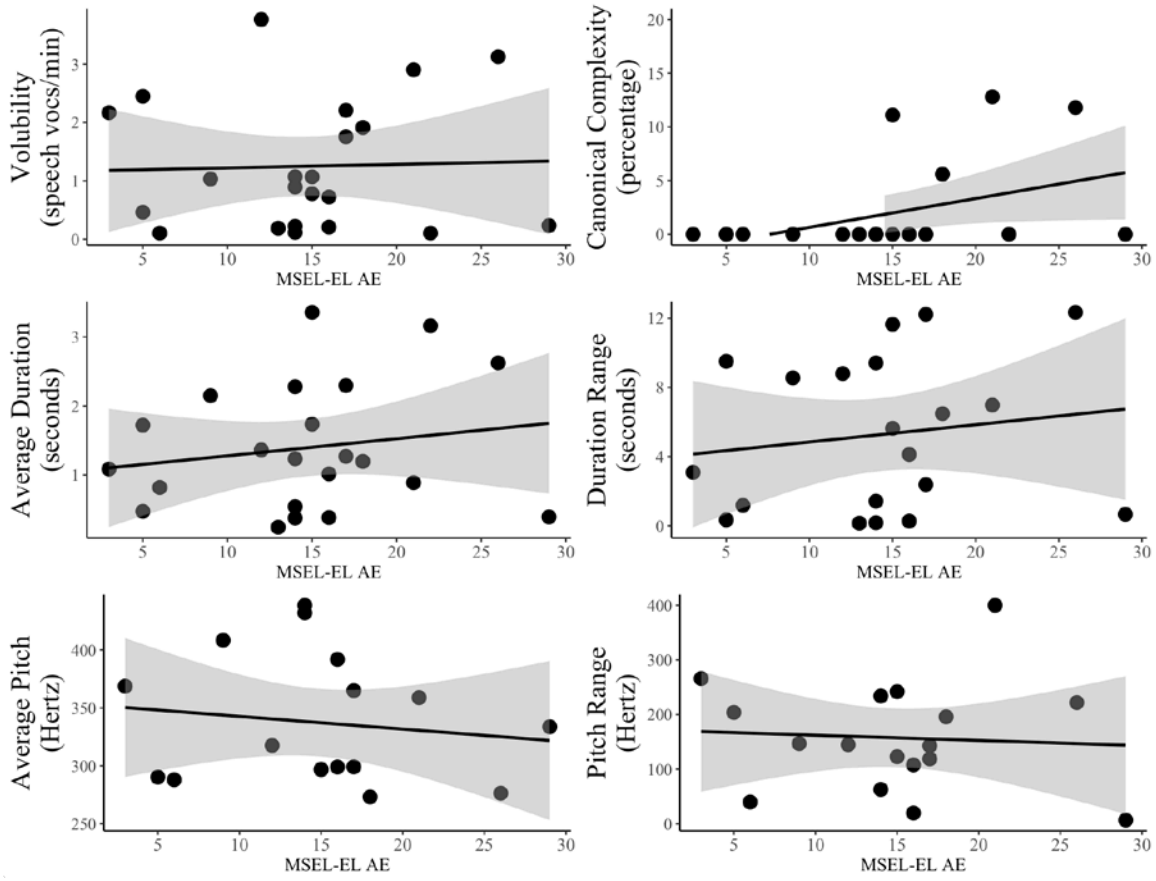


Figure 5. Associations of vocalization features with 24-month Mullen Expressive Language age equivalents in FXS.

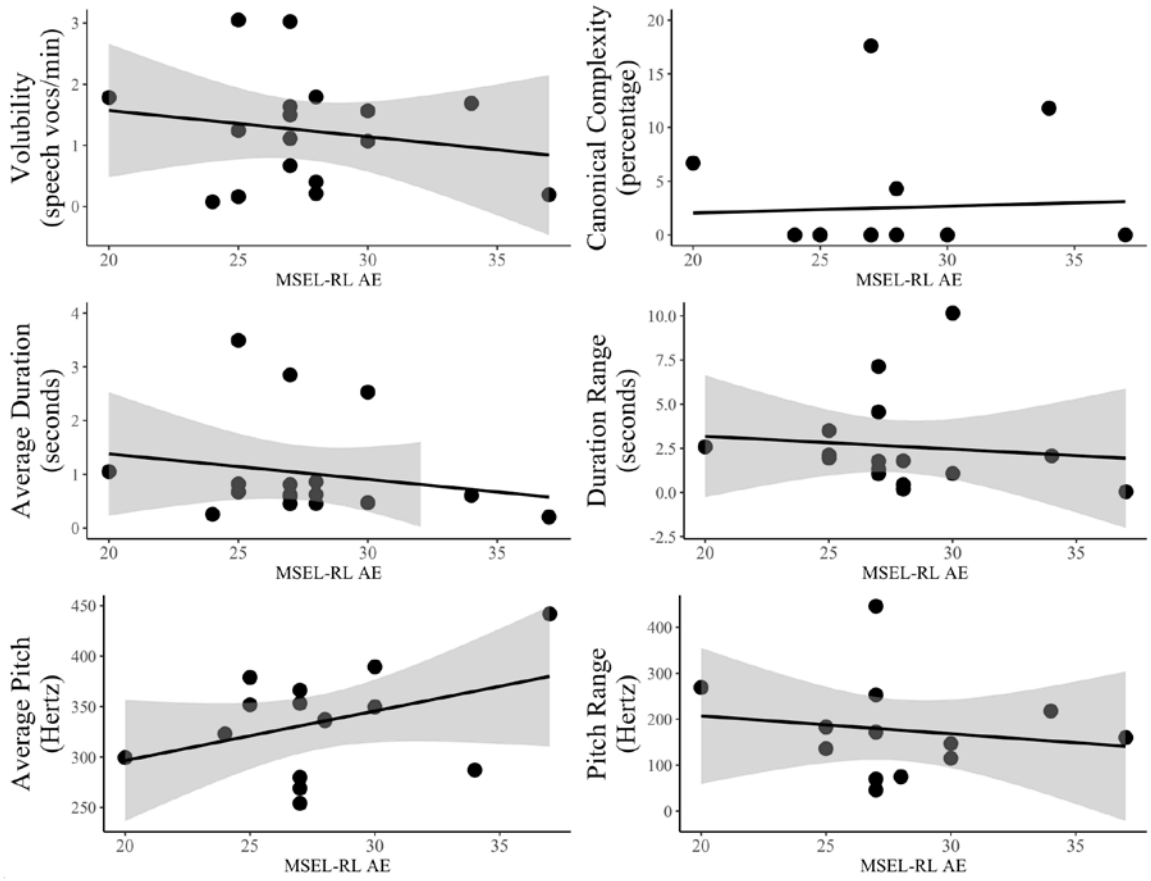


Figure 6. Associations of vocalization features with 24-month Mullen Receptive Language age equivalents in TD.

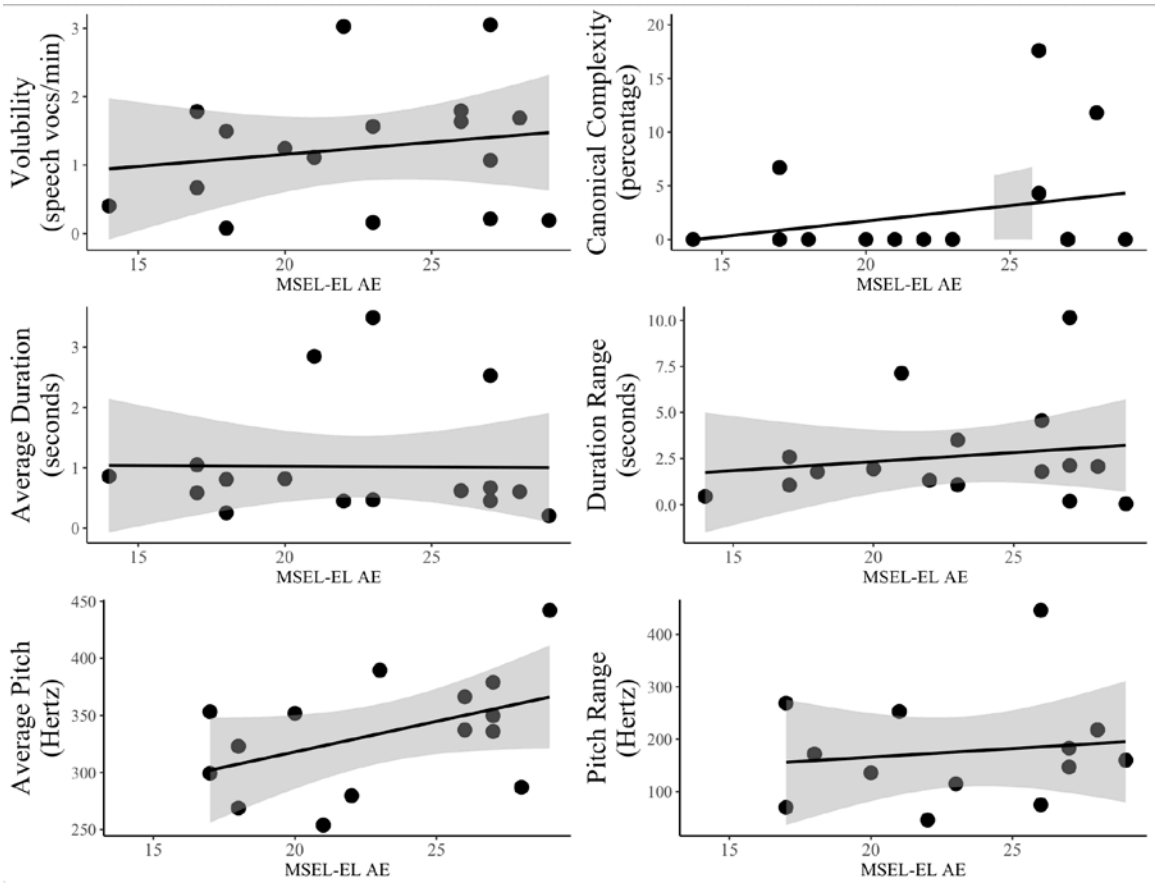


Figure 7. Associations of vocalization features with 24-month Mullen Expressive Language age equivalents in TD.

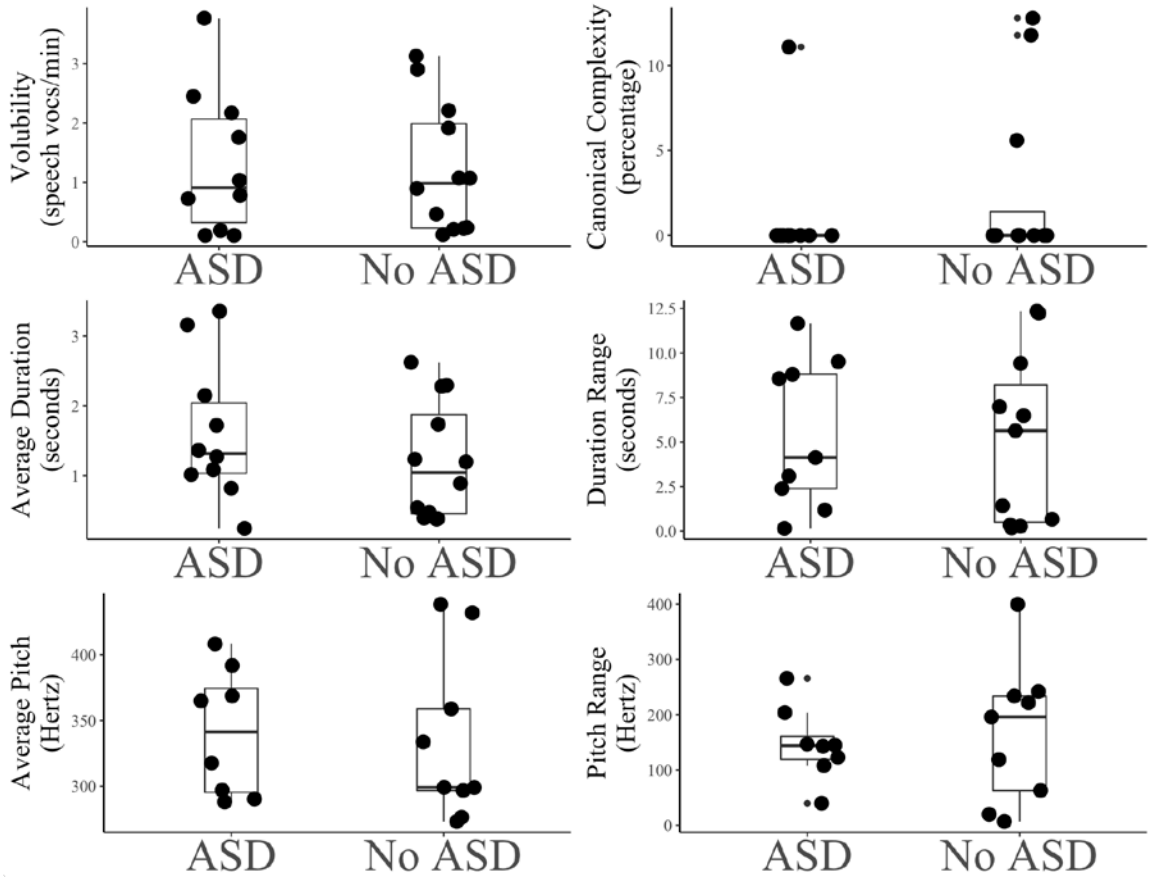


Figure 8. Vocalization features predicting 24-month ADOS-T Social Affect Symptom Severity Scores in FXS.

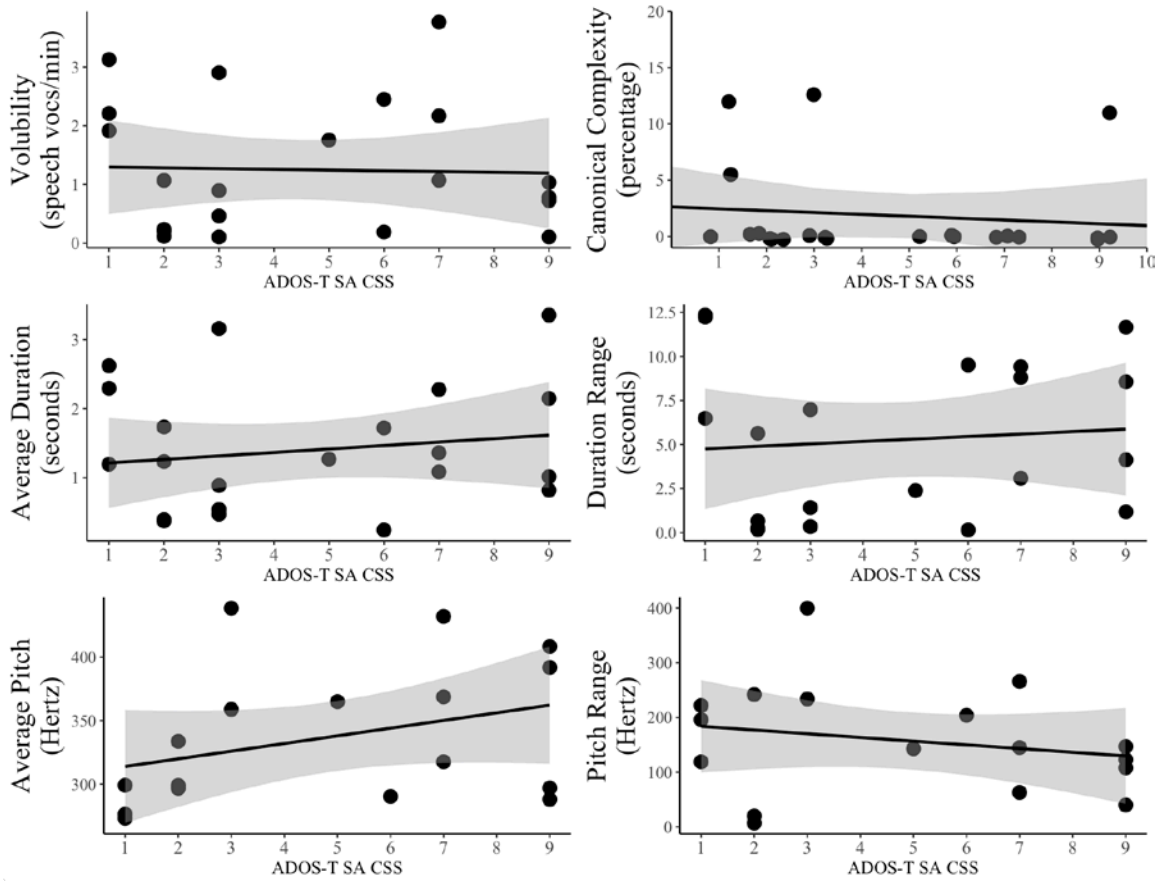


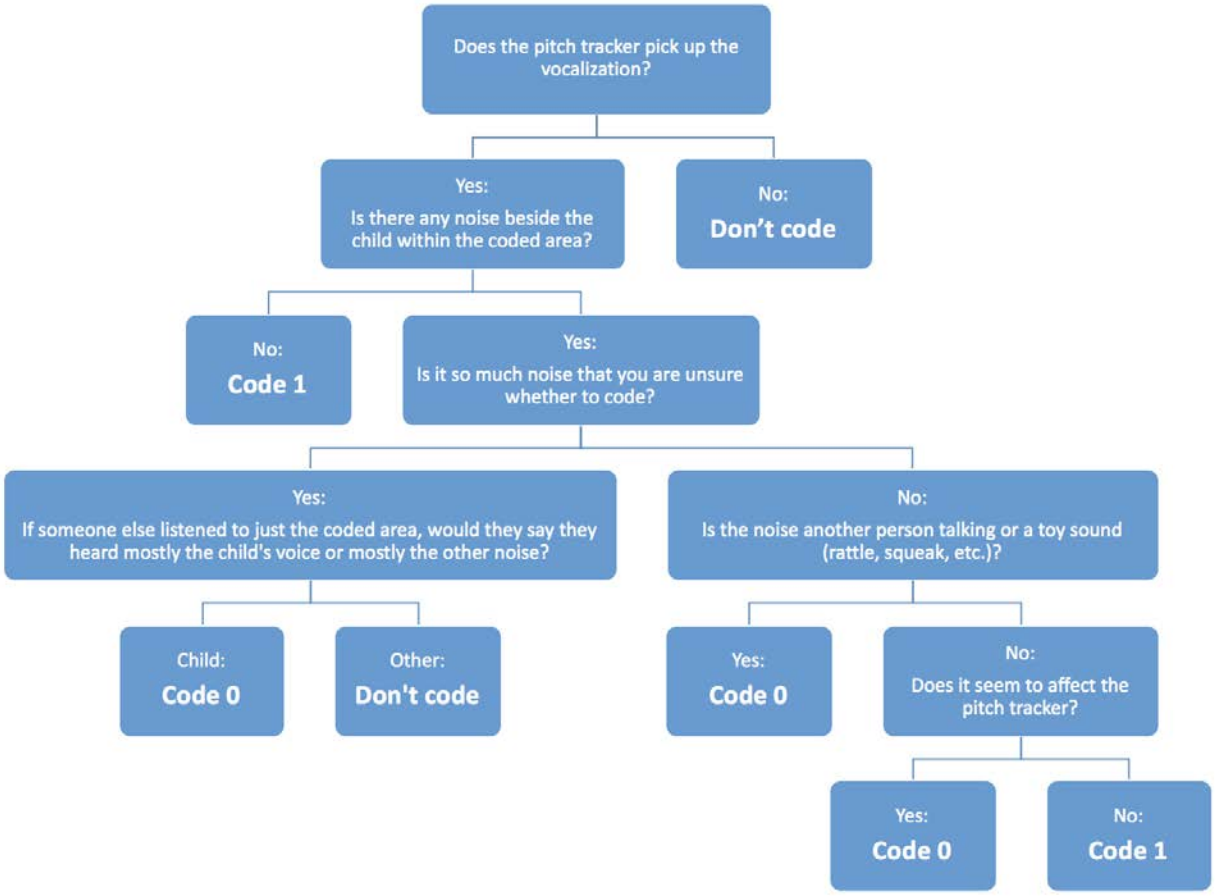
Figure 9. Associations of vocalization features with 24-month ADOS-T Social Affect Symptom Severity Scores in FXS.

APPENDIX C

Excerpts from the Acoustic Coding Manual

Tier Structure

Deciding which vocalizations to code: Refer to the flowchart for each vocalization to determine if and how the vocalization should be coded. Common types of ambiguous codes are described in more detail below.



Do not code:

- **Low amplitude vocalizations**. A vocalization is considered low amplitude if there is not sufficient information in the spectrogram, waveform, or the pitch tracker to determine where the vocalization begins or ends.
- **Noisy vocalizations**. A vocalization is considered noisy if the coder cannot determine whether the sound is the child vocalizing (e.g., the sound could have been made by another person or a noise from the background) or there is so much

background noise that the coder cannot clearly determine what syllable type the child used.

- **Non-speech**. A vocalization is considered non-speech (i.e., not a canonical or precanonical vocalization) if it does not fit precisely in any of the non-speech categories (i.e., laugh, cry, grunt, squeal, bodily function sound).

You may code:

- ***Some* background noise**. If there is noise overlapping any part of the child's vocalization (i.e., beginning, during, or end), but the coder can make a fairly certain determination of what syllable type the child has used.

Tier 1: Codability

There are various ways that background noise (e.g., toy noises, examiner or mother talking during assessment, TV/dog/other child in background) can impact the quality of the acoustic analyses we will be conducting. First, background noise adds another sound to the sound of the child's vocalization. This means that when we pull out acoustic properties of the vocalization, such as the pitch, this may not be an accurate representation of the true properties of the child's vocalization due to the additional sounds from the background noise. Second, background noise that occurs at the beginning or end of the child's vocalization can make it difficult to tell where the child truly begins or stops vocalizing. Again, this may obscure duration of vocalization.

Codability codes:

Category	**CODE**
Codable: No background noise • Coded area contains ONLY the sound of the child vocalizing, or if there are other noises within the boundaries, they do not impact the pitch tracking of the vocalization. • Boundaries of the vocalization are clearly/easily established.	1
Not Codable: Some background noise • Coded area contains noise that OVERLAPS one or both of the boundaries of the vocalization. o Mother/examiner talking o Toy noise (rattle, squeaky toy, bell) o Noise for toys or child banging on table • Coded area contains noise other than the child WITHIN the code boundaries that affects the pitch tracking for the vocalization.	0

Tier 2: Speech Type

Speech Type codes:

Category	Examples	CODE
Vegetative sounds	Cough, sneeze, burp, hiccup, raspberries, sniffling	DO NOT CODE
Non-speech	Cry, whine, growl, grunt, squeal, screech, laugh	0
Pre-canonical/Emerging canonical	ah (or any vowel sound), wa, ya, mm (or any single consonant sound without an accompanying vowel sound) Any kind of mumbling sound (lots of different consonant and/or vowel sounds mashed together) Any vocalization with a single elongated sound (e.g. "mmmmmmmah" would be precanonical, while "ma" would be canonical)	1
Canonical	Consonant-Vowel (CV) sounds: ba, da, ga, ma, na, ta, ka, pa (rare), sa (rare), ha (rare), la (rare) Consonant-Vowel-Consonant (CVC) sounds: "mab", "gad", "dap", etc.	2

Note: All codes are in lowercase.

Tier 2 Coding Considerations:

- The distinction between codes in *different* categories is VERY important, so if you are having a difficult time coming to a consensus on those codes, use the following procedure:
    - If another coder is in the lab, play the sound clip for them and ask what they hear. Do not tell them ahead of time what codes you are trying to distinguish between – they should hear the vocalization blindly.
        - If they hear one of the codes you were trying to decide between, use what they hear as the deciding vote.
        - If they hear something completely different from what you hear, email your coding supervisor with file name and time of code for review

    - *Note*: Sounds that are ambiguous enough to leave you unsure of whether you are hearing a consonant sound are most likely marginal babbling (i.e., precanonical).
- Drawn out vowels should be coded as one long vocalization, even if they seem to change pitch or the type of vowel sound.
    - *Exception*: If there is a breath or short pause between the vowel sounds, then separate them into two codes.
- For canonical codes, each code should bound either a CV or a CVC vocalization. So, while "ba-ba-ba" would be broken into three codes, "bab" would be a single code.

Coding Procedures

- Coding should ALWAYS be completed while wearing headphones.
- When listening to the file for vocalizations, zoom the viewing area to a 50-second window
    - Note the time when you would like to start viewing
    - Open the "View" menu
    - Select "Zoom…"
    - Start the window at the time noted previously and set it to end 50 seconds later
- When you hear a child vocalization:
    a. Listen until there is a clear 1-2 second break after the vocalization when the child is not speaking (you want to make sure you are hearing the full vocalization or group of vocalizations)
    b. Zoom into section with just the vocalizations (at least a 1-second window, more as needed to be able to see the whole group of vocalizations)
    c. Listen to the vocalization group again and place preliminary boundaries at the start and end of each vocalization in the group
        i. Place the boundary only on Tier 1
        ii. Breaks in vocalization groups should be placed based on the number of syllables you hear in the vocalization.
            1. Listen to the group of vocalizations and determine how many syllables you hear, and break up the group of vocalizations based on this determination.
            2. This may be slightly different for cries/whines, which often have a "eh-heh" sound to them – in this case, the "eh-heh" is one syllable. If there are multiple "eh-heh"s, then breaks should be placed as if "eh-heh" is one syllable.
        iii. This may be guided by the blue pitch lines
        iv. Bouts of non-speech (i.e. squeal, cry, laugh, growl) should be broken up based on when the child takes a breath
        v. For vocalizations with more than one syllable (e.g."ba-da"), each syllable should have two boundaries (i.e. each syllable gets its own beginning and end boundary – if the two syllables do seem like

they are continuous, place the end boundary of the first syllable and the beginning boundary of the second syllable on consecutive upward crosses)

vi. Breaths may or may not be included in the vocalization – if they sound like their own separate "syllable," don't need to code them, but if they are part of one continuous vocalization, include them
    1. *Note*. Breathing out often sounds like "heh" or "uh" – be careful in distinguishing a breath from a "sa" or "ah" code.

2. You'll then work through each vocalization in the group, using the following procedure:
    a. Zoom to 0.15 seconds before and 0.15 seconds after the first boundary and find a more precise placement
        i. If there is a clear change in waveforms, adjust the boundary to that location
            1. If there's not a clear change but it looks like your preliminary boundary might have been off, you can adjust the preliminary boundary to make the next step more straightforward
        ii. If there is not a clear change, determine the edge of the vocalization by listening to various increments around the boundary:
            1. Select a window about 0.01 seconds long before the boundary (or after the boundary if you're working on the end bound)
            2. Keep adding 0.01 seconds (approximately one blue dot in the pitch tracker) on the other side of the boundary
            3. Once you listen to an increment where you think you hear the vocalization, place the boundary on the increment before (so that it includes the edge of the vocalization)
                a. You may have to re-listen to some of the increments to decide where the edge is
            4. If you've rotated through the increments more than three times and they still all sound the same, align the boundary with the blue pitch line
    b. Zoom in right before and after the boundary (approximately one blue dot before and one blue dot after) and adjust the boundary to the closest upward cross (see "Placing Boundaries" section)
    c. Zoom in again and adjust boundary so it is precisely on the upward cross
    d. Repeat steps a-c for the end boundary
    e. When both beginning and end boundaries have been set, zoom to 0.06 before and after the vocalization to determine if there is any background noise.
        i. If there is no sound other than the child in this window, code "1" on Tier 1
        ii. If you hear sound other than the child in this window:

1. If it overlaps the beginning and/or the end boundary, code "0" on Tier 1
2. If it does not overlap a boundary, but does appear to affect the pitch tracker, code "0" on Tier 1
3. Anything else can be coded a "1" on Tier 1

f. Repeat steps a-e for each vocalization in the group
g. When each vocalization in the group has been coded, listen to a selection containing all the vocalizations one last time
    i. *Note*. When breaking up "laugh" and "cry" codes, be aware that these types of vocalizations often form a group, with multiple coded areas making up a "laugh group" or "cry group". There may be some codes in these groups that when listened to individually, sounds like vowel sounds or other types of Syllables. Nevertheless, they should still be coded as "laugh" or "cry," based on whether they seem to fit into the group of other "laugh" or "cry" codes.

3. NOTE. When you finish coding a non-reliability file, create a copy of the textgrid that does not include your initials in the file name.

APPENDIX D

Supplemental Tables

Spearman Rank-Order Correlations of Vocalization Features with 9-Month Nonverbal Mental Age (NVMA), 9-Month Chronological Age (CA), and 9-Month NVMA Controlling for 9-Month CA in the FXS and TD Groups

	FXS (n = 22)			TD (n = 17)		
	9-Month CA	9-Month NVMA	9-Month NVMA controlling for 9-Month CA	9-Month CA	9-Month NVMA	9-Month NVMA controlling for 9-Month CA
Volubility	.03	-.02	-.03	-.19	-.34	-.32
Canonical Complexity	-.18	.16	.19	.22	.02	-.02
Average Duration	.13	.12	.10	-.09	-.40	-.39
Duration Range[a]	.18	.07	.02	-.27	-.26	-.22
Average Pitch[b]	.16	.18	.16	.45	.55[†c]	.51[†c]
Pitch Range[b]	.17	-.03	-.05	.19	-.24	-.29

Note. [†]$p < .10$, *$p < .05$, **$p < .01$, ***$p < .001$. [a]Samples used for correlations involving duration range are smaller (FXS n = 20; TD n = 16) due to some participants having too few eligible vocalizations (<2) to calculate a range. [b]Samples used for correlations involving average pitch and pitch range are smaller (FXS n = 17; TD n = 13) due to some participants having 1 or fewer vocalizations with Codability = 1. [c]Value becomes non-significant after correcting for multiple comparisons using the Holm-Bonferroni Sequential Correction.

Spearman Rank-Order Correlations of Vocalization Features with 24-Month Nonverbal Mental Age (NVMA), 24-Month Chronological Age (CA), and 24-Month NVMA Controlling for 24-Month CA in the FXS and TD Groups

	FXS (*n* = 22)		TD (*n* = 17)	
	24-Month NVMA	24-Month NVMA controlling for 24-Month CA	24-Month NVMA	24-Month NVMA controlling for 24-Month CA
Volubility	-.02	-.11	-.38	-.41
Canonical Complexity	.44*[c]	.41[†c]	-.25	-.24
Average Duration	.13	.09	-.20	-.24
Duration Range[a]	.11	.06	-.35	-.28
Average Pitch[b]	-.13	-.08	-.20	.04
Pitch Range[b]	.19	.12	-.39	-.32

Note. [†]$p < .10$, *$p < .05$, **$p < .01$, ***$p < .001$. [a]Samples used for correlations involving duration range are smaller (FXS $n = 20$; TD $n = 16$) due to some participants having too few eligible vocalizations (<2) to calculate a range. [b]Samples used for correlations involving average pitch and pitch range are smaller (FXS $n = 17$; TD $n = 13$) due to some participants having 1 or fewer vocalizations with Codability = 1. [c]Value becomes non-significant after correcting for multiple comparisons using the Holm-Bonferroni Sequential Correction

Spearman Rank-Order Correlations of Vocalization Features and Semi-Partial Correlations of Vocalization Features Controlling for 9-Month Nonverbal Mental Age and 9-Month Chronological Age in FXS and TD Groups

No Controls	1	2	3	4[a]	5[b]	6[b]
1. Volubility	—	.41[†c]	.40[†c]	.77***	-.14	.62**[c]
2. Canonical Complexity	.46[†c]	—	.27	.46*[c]	-.41	.38
3. Average Duration	.05	.10	—	.93***	-.18	.18
4. Duration Range[a]	.27	.24	.71**	—	-.22	.27
5. Average Pitch[b]	-.19	-.35	-.47	-.34	—	.00
6. Pitch Range[b]	.07	.52[†c]	.45	.64*[c]	-.12	—
Controlling 9-month CA and 9-month NVMA	1	2	3	4[a]	5[b]	6[b]
1. Volubility	—	.41**[c]	.41**[c]	.77***	-.11	.64***
2. Canonical Complexity	.51**	—	.27[†c]	.46**	-.41*[c]	.38*[c]
3. Average Duration	-.03	.13	—	.93***	-.15	.20
4. Duration Range[a]	.21	.29	.69***	—	-.20	.29
5. Average Pitch[b]	-.04	-.46*[c]	-.43*[c]	-.25	—	-.02
6. Pitch Range[b]	.03	.55**[c]	.44*[c]	.64**	-.07	—

Note. [†]$p < .10$, *$p < .05$, **$p < .01$, ***$p < .001$. Spearman correlations for FXS ($n = 22$) are presented above the diagonal, and Spearman correlations for TD ($n = 17$) are presented below the diagonal. [a]Samples used for correlations involving duration range are smaller (FXS $n = 20$; TD $n = 16$) due to some participants having too few eligible vocalizations (<2) to calculate a range. [b]Samples used for correlations involving average pitch and pitch range are smaller (FXS $n = 17$; TD $n = 13$) due to some participants having 1 or fewer vocalizations with Codability = 1. [c]Value becomes non-significant after correcting for multiple comparisons using the Holm-Bonferroni Sequential Correction.