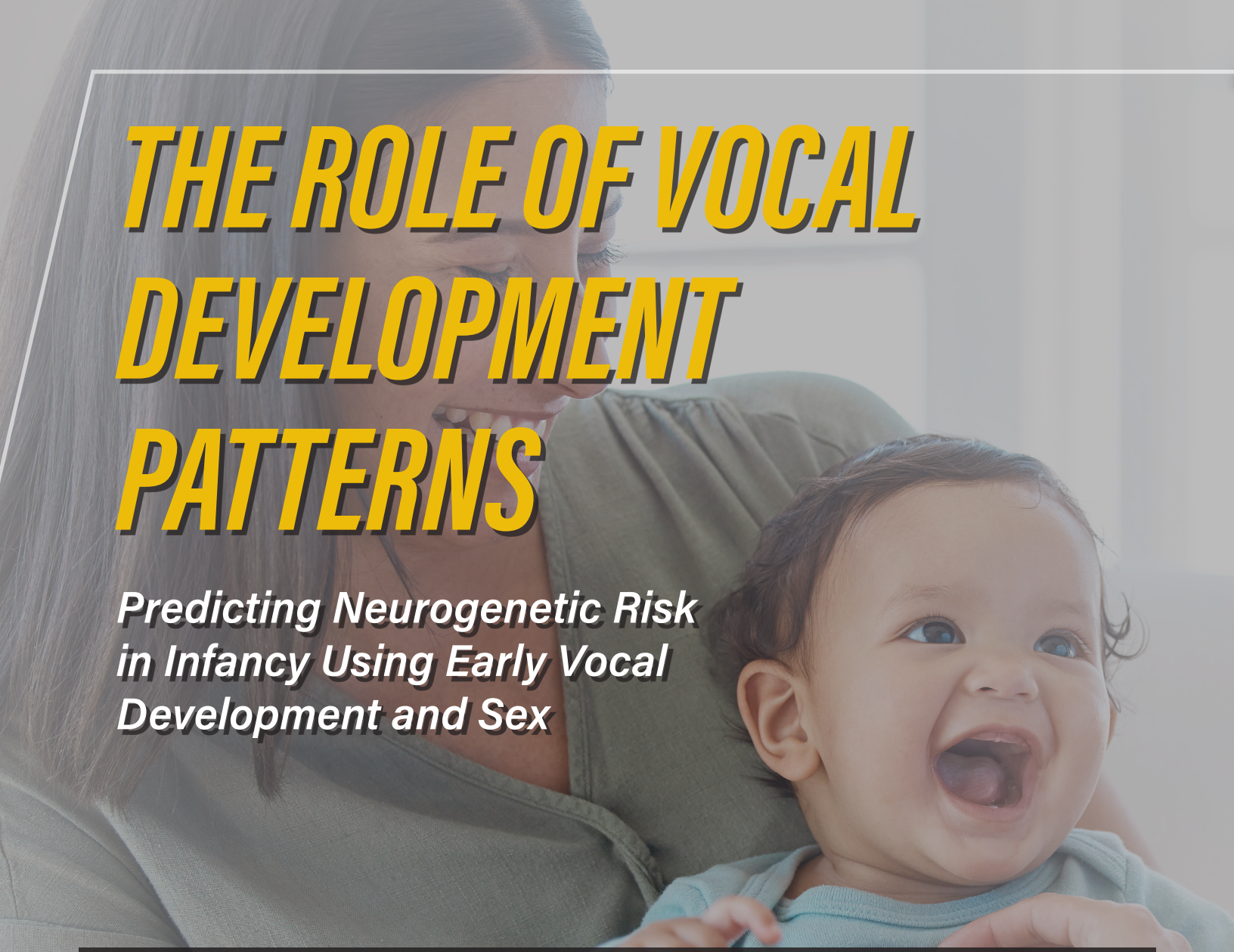

THE ROLE OF VOCAL DEVELOPMENT PATTERNS

Predicting Neurogenetic Risk in Infancy Using Early Vocal Development and Sex

Abstract

Extant literature documents a higher rate of language/speech disorders in males; however, despite sex being a potential moderator of outcomes, we do not know what role it plays in early vocal behavior of infants at high risk for such disorders. The purpose of this study was to ask: (1) Do high-risk infants demonstrate atypical vocal development patterns? (2) Is the quality and quantity of early babble distinct for male and female infants, and does this pattern vary across risk? To answer these questions, we examined the canonical babbling ratio (CBR; the ratio of canonical syllables, those with a consonant and vowel, like "ba," to all babbled syllables) and number of babbled syllables in data collected from 89 high- and low-risk 6- to 18-month-olds. The infants were divided into four groups: infants with Angelman, Down, and fragile X syndromes, and infants at low risk for speech and language disorders. Each participant was recorded for one day using a digital recorder. After recording, speech produced by the infant was extracted and annotated by 3 trained undergraduate coders for the number of canonical and other syllables produced. We ran ANOVAs to explore group and sex differences, which revealed a main effect of risk group, but no main effects or interactions of sex with our dependent variables (number of syllables, CBR). Thus, results revealed group differences, but not effects of sex, suggesting that sex does not relate to risk in vocal production, which could contribute to improved early diagnosis of speech and language disorders.

Keywords

infants, babbling, canonical babble, syllables, neurogenetic syndromes, vocalizations

https://doi.org/10.7771/2158-4052.1605

Student Authors

Back row from left to right: **Victoria Tuell**, **Conghao Gao**, **Athena Fordwor**, **Madison Chin**, **Kaylee Bobay**, **Erin Lee**

Front row from left to right: **Rachel Siela**, **Tiernan McDivitt**, **Laraib Kausar**, **Laurel Williams**

Not pictured: **Alyssa Cregg**, **Deeksha Handa**, **Grace Strabala**

All authors are undergraduate research assistants in the Purdue Infant Speech Lab, with a variety of majors including Speech, Language, and Hearing Sciences, Psychology, and Brain and Behavioral Sciences and ranging from sophomores to graduating seniors. They began working on their research publication in the fall of 2022 and presented their completed work at the Purdue Undergraduate Research Conference in the spring of 2023. All graduating seniors—Rachel Siela, Kaylee Bobay, Erin Lee, and Tiernan McDivitt—will be attending graduate programs to pursue master's degrees in speech-language pathology.

Mentors

Dr. Amanda Seidl is a professor in the Department of Speech, Language, and Hearing Sciences. The overarching goal of her research program is to discover how language comes to the child. In past work, she has focused heavily on how the child learns their phonological system. In her current research program, she is exploring early predictors of language outcomes. Specifically, this work investigates the ways in which measures of early speech perception, production, and the input to the child relate to later language in both typical development and in children at risk for autism spectrum disorders.

Dr. Bridgette Kelleher is an associate professor of psychological sciences and serves as Director of the Purdue Autism Research Center. She obtained her PhD from the University of South Carolina (2015) after completing her APA-certified clinical residency at the Medical University of South Carolina/Charleston Consortium. Kelleher's internationally recognized research program focuses on optimizing clinical outcomes in "high risk" populations, with a particular focus on using technology-driven solutions to assess and treat families affected by rare disorders.

Erica Hill is a graduate student researcher for the Infant Speech Lab at Purdue University and is currently pursuing her master's in Speech-Language Pathology. Her research and clinical interests include the birth to age 3 populations, autism spectrum disorder, neurodevelopmental disorders, early identification and early intervention, infant vocal development, how play affects language development, and gestalt language processing.

Laraib Kausar is currently pursuing her doctorate of audiology at Purdue University and is also a graduate research assistant in the Infant Speech Lab. She is interested in working with the pediatric population and is excited to learn more about cochlear implants, diagnostic and early hearing detection and intervention, and educational audiology.

Lisa Hamrick is a PhD student in the Clinical Psychology program at Purdue University. Hamrick joined Dr. Kelleher's Neurodevelopmental Family Lab in 2016, where her research focuses on early language and social communication development of infants and children at risk for autism spectrum disorder. Hamrick was awarded a National Institutes of Health predoctoral fellowship in 2018 to examine the association of early social communication features and autism symptoms in preschoolers with neurogenetic syndromes, and her dissertation was a downward extension of this research to explore these associations in infants and toddlers with neurogenetic syndromes. Hamrick is currently in the final year of her doctoral degree, completing a predoctoral internship at the Medical University of South Carolina before graduating in August.

A long literature documents a relationship between sex and both the development of linguistic skill and the prevalence of language disorders. For example, Bornstein, Hahn, and Haynes (2004) examined male and female children aged from 1 to 7 years via questionnaires and interviews with participants' parents, standardized assessments in a laboratory setting, and transcripts of the children's spontaneous speech. While results revealed that complete and heterotypic stabilities were equal in female and male children, overall, female children tended to outperform male peers on a variety of linguistic skills. Similarly, Kaushanskaya, Gross, and Buac (2013) examined children learning either phonologically familiar or unfamiliar words (with either a familiar or unfamiliar referent) and found that female children tended to outperform male children on word-learning tasks. Such sex-related differences in language learning could contribute to explaining why sex is a known risk factor for language disorders or delays. For example, in a large UK study on over 6 million children, Lindsay and Strand (2016) found that male children show a much higher prevalence (155% increased risk) of speech, language, and communication needs as compared to female children. Similarly, diagnosis of developmental language disorder has been found to be more prevalent in males than females (2–3 times more likely; Chilosi et al., 2021). Understanding the association between sex and both specific speech and language disorders and language development for children is important since language delays and disorders can contribute to later negative outcomes in both social and educational settings.

Despite sex differences being well established in the literature, questions remain about when and how such sex differences in language skill emerge. Examining early infant vocalizations could tell us about whether sex-related differences in language are present at the very earliest stages of production—for example, whether sex impacts the production of consonant-vowel or canonical syllables (e.g., syllables like "ba"). Cychosz et al. (2021) examined the development of babble and sex in infants growing up in a diverse set of cultures and languages. Results revealed no significant sex differences in the ratio of canonical babbling; however, there may be other sex-related differences in early vocal production. For example, Oller et al. (2020) examined protophone production (precursors to speech that include vowel-like

sounds, squeals, and grunts) by recording 35 female and 65 male infants once a month from 0 to 12 months. They found that males had a 24% higher protophone rate than females, but that among males, increased protophone rate did not correlate with a higher canonical babbling ratio. Thus, their results suggest that male infants make sounds more frequently than female infants, but that regardless of this higher volubility, male and female infants may be matched on vocal skill. Notably, Oller et al. (2020) also measured the differences in protophone rate and canonical babbling rate for infants who were considered high risk (because they had an older sibling diagnosed with autism) and found specifically that the high-risk *males* had the highest protophone rate, suggesting that risk and sex may interact in early vocal behaviors.

Why might these differences emerge? To understand potential biological mechanisms of these associations, Quast et al. (2016) examined whether sex effects on babble were related to hormone differences in male versus female infants. Specifically, they gathered estradiol and testosterone concentration levels in postnatal infants going through the mini-puberty (at 4 and 20 weeks of age). Results revealed a positive association between vocalization skill and estradiol and a negative association between vocalization skill and testosterone, which suggests that hormone concentrations may underlie any sex-related differences in infant vocal behaviors.

Another biological difference that is known to moderate language outcomes is the presence of genetic diagnoses associated with atypical development. A number of neurogenetic syndromes have been associated with atypical language outcomes, including Angelman and fragile X syndromes. Although delayed or absent speech is present in a number of these conditions, it is unclear whether severity of outcomes varies by sex, similar to outcomes in non-neurogenetic groups. Interrogating this possibility may help elucidate the ways in which genetics and sex-related differences intersect. Given the higher prevalence for speech and language disorders in males than females, we expect that there may be a relationship between sex-related differences in early language development, including among children at neurogenetic risk for atypical outcomes. Following from this, we explored whether babbling differences in high-risk populations emerge early in linguistic development and whether they are more marked in male than female infants. We asked two specific questions:

1. Do infants with neurogenetic disorders demonstrate atypical vocal development patterns in both the production of canonical syllables and the number of syllables produced?
2. Are patterns of early canonical babble and the number of babbled syllables distinct for male and female infants across neurogenetic groups?

METHODS

Participants

We included 89 infants between the ages of 6 months and 18 months (spanning one year of vocal development) from a currently funded project in our labs (see acknowledgments for the specific grants supporting this work). Infants were classified as "high risk" if they had a clinical diagnosis of a neurogenetic disorder such as Angelman syndrome (AS), fragile X syndrome (FXS), or Down syndrome (DS), and were classified as "low risk" (LR) if they were typically developing as per caregiver report and had no known family history (first-degree relative) of a speech or language disorder (see Table 1 for demographic information).

Data Collection

Participants were recorded over the course of one full day using the Language ENvironment Analysis Digital Language Processor (LENA®; Xu, Richards, & Gilkerson, 2014), a recorder designed to be worn inside the breast pocket of an infant vest (see Figure 1). Segments that included infant vocalizations were selected via algorithms from LENA's software. Using our own scripts, identical to those used in Semenzin et al. (2021) and Hamrick et al. (in press), we selected a subset of 30 5-min segments (10 were collected from periods of the highest volubility for the infant and 20 were randomly selected from the recording by a script). Infant babble was then annotated by humans (see below) for each infant within these selected segments. In addition to collecting the infant vocalizations, a Vineland-3 (Sparrow et al., 2021) was gathered from each participant. We used this standardized assessment to calculate an adaptive behavior composite (ABC), which is

TABLE 1. Participant table showing number of infants in each clinical group, age, and Vineland scores for infants within each group and sex.

Clinical Group	Age	Risk Group	Sex	N	Vineland ABC
Angelman Syndrome	M = 14.2 (SD = 4.2)	High	Female	6	M = 71.4 (SD = 16.9)
			Male	2	M = 62.0 (SD = 22.6)
Fragile X Syndrome	M = 15.0 (SD = 2.5)	High	Female	5	M = 87.3 (SD = 17.8)
			Male	3	M = 78.0 (SD = 7.1)
Down Syndrome	M = 12.8 (SD = 3.9)	High	Female	12	M = 88.5 (SD = 14.0)
			Male	18	M = 74.1 (SD = 8.2)
Low Risk	M = 11.4 (SD = 3.2)	Low	Female	23	M = 103.8 (SD = 14.4)
			Male	20	M = 101.9 (SD = 10.7)

FIGURE 1. LENA device inside the infant vest used to record infant vocalizations.

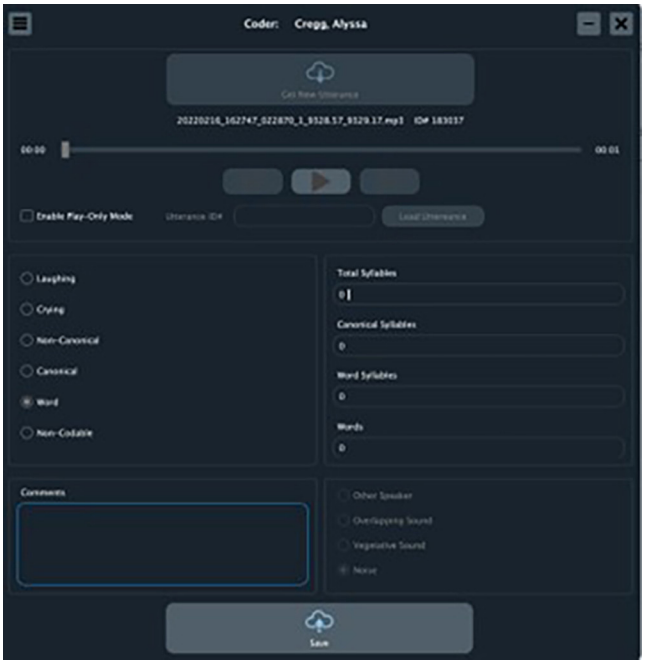

FIGURE 2. The Vocal Maturity Coding Program interface used to annotate the audio files.

typically low in those who have developmental delays, as can be seen in Table 1.

Human Annotation

Undergraduate students in labs in the Speech, Language, and Hearing Sciences or Psychological Sciences departments were trained on how to classify the vocal maturity of infant vocalizations. After this, they were trained to use a computer interface to annotate the vocal maturity and number of syllables of infant vocalizations using utterances extracted from LENA audio files using the interface in Figure 2.

Undergraduate research assistants annotated each utterance, choosing between six different annotation categories: crying, laughing, noncanonical, canonical, word, or noncodable (Figure 2). They also determined the total number of syllables in each utterance and the number of canonical and noncanonical syllables. Syllables within an utterance were identified as canonical if they contained a consonant and vowel (e.g., "ba"), while a noncanonical syllable contained only vowels or only consonants such as syllabic nasals (e.g., "oh"/"mmm"). Recordings were identified as "noncodable" if they had overlapping speech/sounds (e.g., mother talking over

infant), vegetative sounds (e.g., burps), noise (e.g., loud toys), or were produced by another speaker (e.g., the mother or an older child). Each utterance was coded by three different annotators and was used in data analyses only when 3 of the 3 annotators agreed on the annotation. This "clean" data set resulted in 15,862 utterances containing canonical syllables and 55,867 utterances containing only noncanonical syllables, for a total of 71,729 utterances annotated.

RESULTS

Using the clean data set (with 3/3 annotator agreement), we calculated a canonical babbling ratio (CBR; number of canonical syllables to the number of total syllables [canonical + noncanonical]) for each child and the total number of syllables produced (measured as the sum of canonical and noncanonical syllables) for each child. We then examined whether these two measures varied with risk status and sex (male, female) and whether sex interacted with risk (high-risk [FXS, AS, DS]; LR). Our prediction was that (1) the LR group would present with a higher CBR and greater number of syllables produced, compared to the high-risk group, and (2) that high-risk male infants would be more affected overall than female infants.

To test our predictions, we ran ANOVAs with CBR and the number of syllables produced as dependent measures and risk (LR, AS, DS, FXS) and sex as independent measures. Results for the ANOVA on CBR revealed a main effect of risk ($F(3,81) = 3.88$, $p < .012$). There was no main effect of sex ($p = .192$) or interaction between sex and risk ($p = .496$). Follow-up *t*-tests comparing risk groups revealed that CBR was significantly different for LR vs. AS ($p < .005$) and LR vs. DS ($p < .011$), but not LR vs. FXS ($p = .232$). All other differences, including among syndrome groups, were not significant (see Figure 3).

Results for an ANOVA with the same structure, but with number of babbled syllables as the dependent measure, similarly revealed a main effect of risk ($F(3,81) = 3.34$, $p < .023$), and no main effect of sex ($p = .367$) or interaction of sex and risk ($p = .97$). Follow-up *t*-tests comparing risk groups revealed that the number of syllables babbled was significantly different or marginally different for all risk groups as compared with the LR group (see Figure 4; LR vs. FXS ($p < .008$), LR vs. DS ($p < .037$), LR vs. AS ($p < .086$)).

DISCUSSION

Our results revealed that, counter to our predictions, neither CBR nor the number of babbled syllables produced differed between male and female infants, nor did sex interact with risk. However, while sex did not impact CBR or the number of babbled syllables

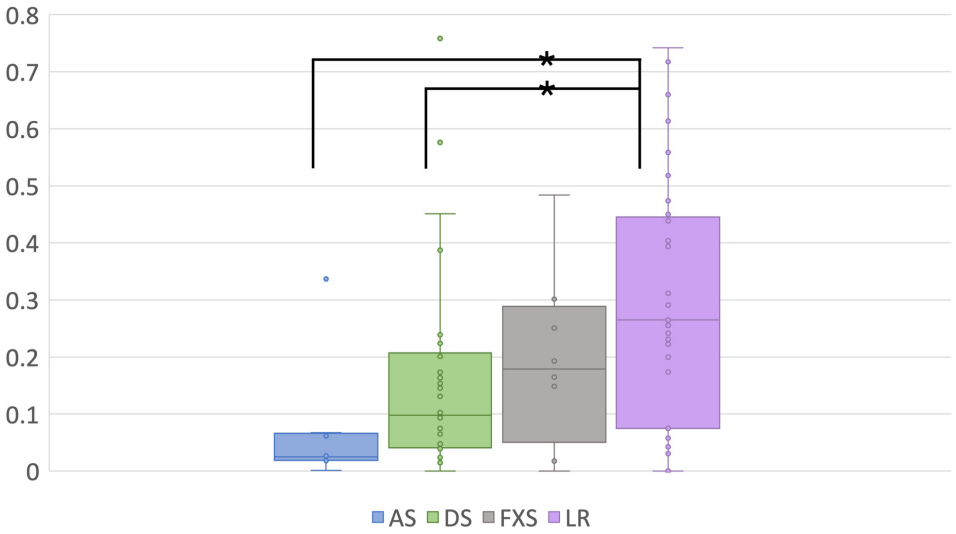

FIGURE 3. CBR by risk group. Groups with asterisks are significantly different.

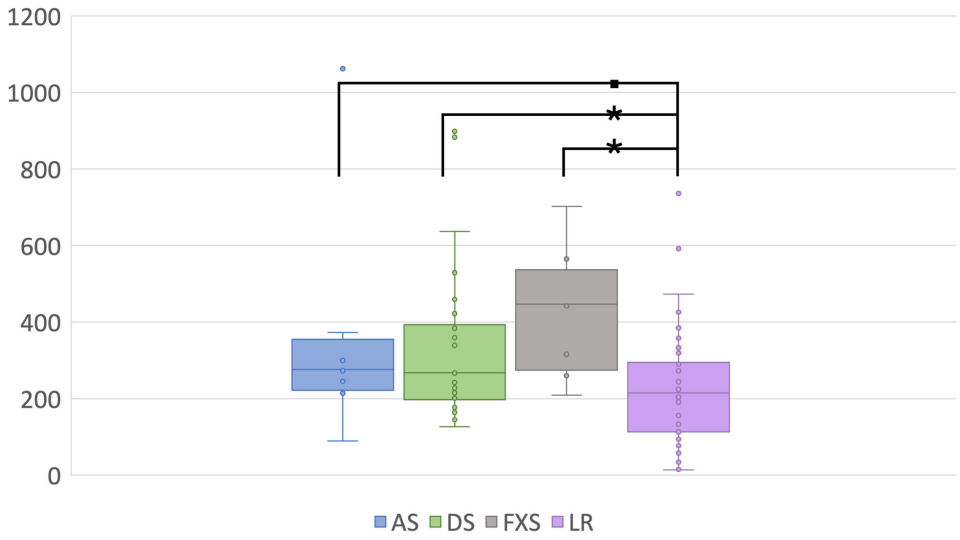

FIGURE 4. Number of syllables produced by risk group. Groups with asterisks are significantly different. Groups with a dot are marginally different.

produced, risk had a clear effect for both measures. Specifically, CBR was higher for LR compared to DS/AS (though not significantly different for LR vs. FXS), and the number of syllables produced was lower for LR compared to FXS/DS (although only marginally different compared to AS). In sum, while CBR and number of syllables produced were not able to differentiate all risk groups in isolation, our results suggested that if they are used in conjunction there may be a better ability to predict risk and to differentiate these risk groups from LR children.[1]

One possible explanation for this pattern of results in which high-risk infants have higher volubility and lower CBR than low-risk infants is that infants in the high-risk groups may be less socially selective in how they vocalize than infants in the low-risk groups. This explanation is motivated by findings from two recent studies. The first shows that low-risk infants display social sensitivity in their babbling behaviors. Specifically, Long et al. (2022) explore infant vocalizations within turn-taking and vocal play (infant alone) contexts and find that during vocal play infants produce more noncanonical syllables and more speech overall, while during turn-taking infants produce more canonical syllables, but fewer syllables overall. Thus, they argue that LR infants are endogenously driven to vocalize (which they display when alone), but engage in more advanced vocalizations in social contexts. Because we might predict that high-risk infants, many of whom had neurogenetic syndromes associated with autistic features, may be less sensitive to social contexts, they may be less able to modulate their vocal behavior as a result of social interaction and may overall display more vocal play–type patterns than turn-taking patterns seen in LR infants. Data to support this hypothesis comes from another recent paper (Swanson et al., 2018) which shows that a subset of 9-month-old infants who are at high risk for ASD show clear patterns of hypervolubility (more vocalizations than is typical). Through the use of LENA recordings, they analyzed groups of low-risk and high-risk 9-month-olds. A subset of these infants showed vocalizations that were 2 standard deviations greater than that in the other infant groups (90% more vocalizations than the low-risk group), but not a greater number of conversational turns (i.e., much of their vocalizations were nonsocial in nature). As a result of this greater vocalization rate (potentially more vocalizations without an adult having an opportunity or ability to respond), these children experienced reduced responsivity from the adult interacting with them. This lack of social feedback could impact their CBR if the infant is using social responsivity (i.e., caregivers' selective responses to infants' speechlike vs. nonspeechlike vocalizations) to guide language growth as has been suggested (Goldstein & Schwade, 2008; Warlaumont et al., 2014). Thus, findings from Swanson et al. (2018) parallel some of our findings in which we found that infants with FXS produced the greatest number of syllables, which could suggest that the

hypervolubility we found in infants with FXS is associated with social sensitivity, as individuals with FXS show a high prevalence of ASD outcomes (Kaufmann et al., 2017) and thus likely less social sensitivity. Note in addition that the CBR of infants with FXS was not statistically significantly different from the CBR infants with LR, which suggests that perhaps social feedback may not impact growth of vocal maturity within these infants at greater risk for ASD.

As mentioned, contrary to our hypotheses we did not find a significant effect of infant sex or interactions with sex. However, while sex differences were not shown to have a significant effect on the canonical babbling ratio and the number of babbled syllables in this data, this result could stem from limitations including our small sample size, a broad age range (infants at different developmental stages when sex only potentially shows up as a factor at a particular point in development), and variability in the data set (lots of different diagnoses). Therefore, while it is still possible that sex-related language effects are present in infancy, our results suggest that, if present, sex effects are either weak or not present within the specific age range explored. Only future exploration with more age ranges and larger data sets will help adjudicate between these alternatives. Nonetheless, since sex does not appear to be a factor that interacts with risk, based on this data set of 89 high- and low-risk infants between ages 6 and 18 months, it makes the task of using CBR and number of babbled syllables a potentially simpler diagnostic tool since these measures would not have to be normalized to each sex to provide diagnostic strength within this age window. However, exploration of this question with larger and distinct clinical samples may be necessary before concluding that sex does not interact in any way with the variables in question.

Critically, our results regarding the ability to use CBR and number of babbled syllables to predict neurogenetic group membership contribute to fill an important gap in the literature as well as provide a potentially powerful early screening measure for later diagnosed disorders. Specifically, some of the neurogenetic disorders explored in this work are diagnosed quite late in development (e.g., not until 3–5 years in FXS, AS, and/or Rett syndrome), which can lead to children missing both critical linguistic milestones and opportunities for intervention. Similar to our study, earlier work with different populations suggests that early delay in canonical babble can predict developmental disorders (e.g., Lang et al., 2019; Patten et al., 2014) as well as later language outcomes (Oller et al., 2020). For example, Patten et al. (2014) conducted a retrospective video analysis of vocalizations in children with ASD and typical development and found that children with ASD showed lower CBR at 9–12 months and lower volubility at 9–12 and 15–18 months of age. Further, an analysis that combined both variables (CBR, volubility) was able to predict 75% of ASD classification, suggesting that these might be a good early screening tool for ASD. However, Patten et al.'s study had several limitations. First, the videos explored were short in duration and thus might not be representative of infants' behavior given high variability in infant babbling behaviors—even within the typically developing population. Second, there was no comparison group composed of infants with non-ASD disabilities, which may be problematic given that we cannot assess the specificity of these potential diagnostic measures. Third, the sample was very homogeneous in SES and thus might not provide a reliable diagnostic tool for diverse populations since babble behaviors may vary with settings (e.g., infants growing up in rural areas hear less infant-directed input, which may impact their babbling behaviors (Bergelson et al., 2023). Thus, given these limitations it has been challenging to import such work for use in diagnostic development.

Note, however, that in the current work we address several of the limitations present in this previous work by recruiting a more representative sample (infants were diverse both geographically, socioeconomically, and in their diagnoses), and through collecting daylong recordings so the speech sample obtained was both longer and more ecologically valid. Our results with this larger and more diverse sample support those from this previous body of work showing a role for both number of syllables babbled and CBR in predicting infant risk status and showing an ability to differentiate based on the specific neurogenetic disorder. Further, the data-gathering method used here surmounts several challenges noted in previous work in which child vocalization data was gathered from samples of children with neurogenetic disorders. For example, Lang et al. (2019) suggest that, while some data from child vocalizations exist, it is both minimal (studies involve small numbers of children) and highly variable (even within a diagnostic criteria there are

differences in severity), making it difficult to determine if measures of early vocalizations could help with earlier diagnosis. Further, as a consequence of the rarity of these disorders, for example, ranging from 1 in 700 in DS (Mai et al., 2019) to 1 in 20,000 in AS (Mane & Chatterjee, 2015), with FXS prevalence falling in between 1 in 7,000 males to 1 in 11,000 females (Hunter et al., 2014), finding large sample sizes has been challenging due to both geographic distribution and unequal access to clinical services across the country. This is especially true for those who have low SES, live in rural areas, lack transportation, speak a nonmajority language, lack education, or are a member of a racial or ethnic minority group. Importantly, examining early vocalizations through remotely gathering recordings in the way that we did here may help to expedite diagnosis within these populations and sidestep some of these challenges. Thus, this project demonstrates both the feasibility of this approach as well as the potential diagnostic utility of these remotely gathered vocalization measures.

While the differences in vocal maturity shown in these populations may be a simple result of differences in linguistic ability across risk groups, it is also worthy of mention that alternative explanations other than linguistic ability account for differences in CBR and total syllables between groups. One possible alternative explanation for differences seen in populations of infants explored in this work is that measures of differences in vocal production may be driven by, or interact with, motor skill and motor developmental delays in these clinical populations. For example, children who are at high risk for ASD are more likely to exhibit delays in their gross motor skills in infancy (Iverson et al., 2019). Specifically, LeBarton and Iverson (2016) find that gross motor delays were most prevalent between 5 and 6 months of age with 61% of high-risk infants presenting with a delay at 5 months and 52% at 6 months. These motor delays may be related to vocal maturity. For example, LeBarton and Iverson (2016) show that in this high-risk population the onset of sitting development is significantly correlated with reduplicated babble development and show gestures at 7 months of age. They argue that this correlation may be found because postural changes (rib cage opening, upright head position, trunk control) allow for easier articulation of consonants and vowels. Further, it is also possible that increased motor ability allows for easier manipulation of objects for show gestures and/or because sitting increases opportunities for social interaction, which may lead to more joint attention and eye contact. Our research shows a clear difference in the vocal maturity of infants in different neurogenetic groups; however, future studies should examine the association between vocal development patterns and motor impairments/gross and fine motor development with relation to high-risk groups.

ACKNOWLEDGMENTS

We would like to acknowledge the hard work of our mentors, Dr. Amanda Seidl, Dr. Bridgette Kelleher, Erica Hill, Laraib Kausar, and Lisa Hamrick, for their dedication and guidance. We would also like to congratulate our graduating seniors, Rachel Siela, Kaylee Bobay, Erin Lee, and Tiernan McDivitt. We wish them all the best as they continue their studies into graduate school. Creation of the data set used for this work comes from an F31 award to Lisa Hamrick and a K23 award to Bridgette Kelleher.

Correspondence concerning this article should be addressed to Rachel M. Siela. Email: rmsiela@gmail.com

NOTE

1. To explore whether these two measures could best predict group membership, we constructed a number of regression models. Specifically, we used these models to explore whether CBR, together with the number of babbled syllables, was better able to predict diagnosis than either measure alone. These regression models revealed that the model with both CBR and number of babbled syllables fit the data and had a higher log-likelihood.

REFERENCES

Bergelson, E., Soderstrom, M., Schwarz, I., Rowland, C. F., Ramirez-Esparza, N., Hamrick, L., . . . Cristia, A. (2023). Everyday language input and production in 1001 children from 6 continents. *PNAS*. https://doi.org/10.31234/osf.io/fjr5q

Bornstein, M. H., Hahn, C.-S., & Haynes, O. M. (2004). Specific and general language performance across early childhood: Stability and gender considerations. *First Language*, *24*(3), 267–304. https://doi.org/10.1177/0142723704045681

Chilosi, A. M., Brovedani, P., Cipriani, P., & Casalini, C. (2021). Sex differences in early language delay and in developmental language disorder. *Journal of Neuroscience Research*, *00*, 1–14. https://doi.org/10.1002/jnr.24976

Cychosz, M., Cristia, A., Bergelson, E., Casillas, M., Baudet, G., Warlaumont, A. S., Scaff, C., Yankowitz, L., & Seidl, A. (2021). Vocal development in a large-scale crosslinguistic corpus. *Developmental Science*, *24*(5), e13090. https://doi.org/10.1111/desc.13090

Goldstein, M. H., & Schwade, J. A. (2008). Social feedback to infants' babbling facilitates rapid phonological learning. P*sychological Science*, *19*(5), 515–523. https://doi.org/10.1111/j.1467-9280.2008.02117.x

Hamrick, L. R., Seidl, A., & Kelleher, B. L. (in press). Semi-automatic assessment of vocalization quality for children with and without Angelman syndrome. *American Journal on Intellectual and Developmental Disabilities.*

Hunter, J., Rivero-Arias, O., Angelov, A., Kim, E., Fotheringham, I., & Leal, J. (2014). Epidemiology of fragile X syndrome: A systematic review and meta-analysis. *American Journal of Medical Genetics. Part A*, *164A*(7), 1648–1658. https://doi.org/10.1002/ajmg.a.36511

Iverson, J. M., Shic, F., Wall, C. A., Chawarska, K., Curtin, S., Estes, A., Gardner, J. M., Hutman, T., Landa, R. J., Levin, A. R., Libertus, K., Messinger, D. S., Nelson, C. A., Ozonoff, S., Sacrey, L. R., Sheperd, K., Stone, W. L., Tager-Flusberg, H. B., Wolff, J. J., Yirmiya, N., . . . Young, G. S. (2019). Early motor abilities in infants at heightened versus low risk for ASD: A Baby Siblings Research Consortium (BSRC) study. *Journal of Abnormal Psychology*, *128*(1), 69–80. https://doi.org/10.1037/abn0000390

Kaufmann, W. E., Kidd, S. A., Andrews, H. F., Budimirovic, D. B., Esler, A., Haas-Givler, B., Stackhouse, T., Riley, C., Peacock, G., Sherman, S. L., Brown, W. T., & Berry-Kravis, E. (2017). Autism spectrum disorder in fragile X syndrome: Cooccurring conditions and current treatment. *Pediatrics, 139* (Supplement_3). https://doi.org/10.1542/peds.2016-1159f

Kaushanskaya, M., Gross, M., & Buac, M. (2013). Gender differences in child word learning. *Learning and Individual Differences*, *27,* 82–89. https://doi.org/10.1016/j.lindif.2013.07.002

Lang, S., Bartl-Pokorny, K. D., Pokorny, F. B., Garrido, D., Mani, N., Fox-Boyer, A. V., Zhang, D., & Marschik, P. B. (2019). Canonical babbling: A marker for earlier identification of late detected developmental disorders?. *Current Developmental Disorders Reports*, *6*(3), 111–118. https://doi.org/10.1007/s40474-019-00166-w

LeBarton, E. S., & Iverson, J. M. (2016). Associations between gross motor and communicative development in at-risk infants. *Infant Behavior & Development*, *44,* 59–67. https://doi.org/10.1016/j.infbeh.2016.05.003

Lindsay, G., & Strand, S. (2016). Children with language impairment: Prevalence, associations and ethnic disproportionality in an English population. *Frontiers in Education: Special Educational Needs, 1*. https://doi.org/10.3389/feduc.2016.00002

Long, H. L., Ramsay, G., Griebel, U., Bene, E. R., Bowman, D. D., Burkhardt-Reed, M. M., & Oller, D. K. (2022). Perspectives on the origin of language: Infants vocalize most during independent vocal play but produce their most speech-like vocalizations during turn taking. *PloS One*, *17*(12), e0279395. https://doi.org/10.1371/journal.pone.0279395

Mai, C. T., Isenburg, J. L., Canfield, M. A., Meyer, R. E., Correa, A., Alverson, C. J., Lupo, P. J., Riehle-Colarusso, T., Cho, S. J., Aggarwal, D., Kirby, R. S., & National Birth Defects Prevention Network. (2019). National population-based estimates for major birth defects, 2010–2014. *Birth Defects Research*, *111*(18), 1420–1435. https://doi.org/10.1002/bdr2.1589

Mane, S., & Chatterjee, R. (2015). Angelman syndrome: The blurred lines of interpretation in cognitive defects. *Journal of Pediatric Neurosciences*, *10*(1), 70–72.

Oller, D. K., Griebel, U., Bowman, D. D., Bene, E., Long, H. L., Yoo, H., & Ramsay, G. (2020). Infant boys are more vocal than infant girls. *Current Biology: CB*, *30*(10), R426–R427. https://doi.org/10.1016/j.cub.2020.03.049

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

Quast, A., Hesse, V., Hain, J., Wermke, P., & Wermke, K. (2016). Baby babbling at five months linked to sex hormone levels in early infancy. *Infant Behavior and Development*, *44*, 1–10. https://doi.org/10.1016/j.infbeh.2016.04.002

Semenzin, C., Hamrick, L., Seidl, A., Kelleher, B. L., & Cristia, A. (2021). Describing vocalizations in young children: A big data approach through citizen science annotation. *Journal of Speech, Language, and Hearing Research, 64*(7), 2401–2416. https://doi.org/10.1044/2021_JSLHR-20-00661

Sparrow, S. S., Cicchetti, D. V., & Saulnier, C. A. (2021). *Vineland-3: Vineland Adaptive Behavior Scales*—third edition. Pearson.

Swanson, M. R., Shen, M. D., Wolff, J. J., Boyd, B., Clements, M., Rehg, J., Elison, J. T., Paterson, S., Parish-Morris, J., Chappell, J. C., Hazlett, H. C., Emerson, R. W., Botteron, K., Pandey, J., Schultz, R. T., Dager, S. R., Zwaigenbaum, L., Estes, A. M., Piven, J., & IBIS Network. (2018). Naturalistic language

recordings reveal "hypervocal" infants at high familial risk for autism. *Child Development*, *89*(2), e60–e73. https://doi.org/10.1111/cdev.12777

Warlaumont, A. S., Richards, J. A., Gilkerson, J., & Oller, D. K. (2014). A social feedback loop for speech development and its reduction in autism. *Psychological Science*, *25*, 1314–1324. https://doi.org/10.1177/0956797614531023

Xu, D., Richards, J. A., & Gilkerson, J. (2014). Automated analysis of child phonetic production using naturalistic recordings. *Journal of Speech, Language, and Hearing Research: JSLHR*, *57*(5), 1638–1650. https://doi.org/10.1044/2014_JSLHR-S-13-0037